

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : <b>C12N 15/49, C07K 14/16</b>		<b>A2</b>	(11) International Publication Number: <b>WO 00/39303</b>
			(43) International Publication Date: 6 July 2000 (06.07.00)
(21) International Application Number: PCT/US99/31272		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 30 December 1999 (30.12.99)		<p><b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i></p>	
(30) Priority Data: 60/114,495 31 December 1998 (31.12.98) US 60/156,670 29 September 1999 (29.09.99) US			
(71) Applicant: CHIRON CORPORATION [US/US]; 4560 Horton Street, Emeryville, CA 94608 (US).			
(72) Inventors: BARNETT, Susan; Chiron Corporation, 4560 Horton Street - R440, Emeryville, CA 94608 (US), HARTOG, Karin; Chiron Corporation, 4560 Horton Street - R440, Emeryville, CA 94608 (US), MARTIN, Eric; Chiron Corporation, 4560 Horton Street - R440, Emeryville, CA 94608 (US).			
(74) Agents: DOLLARD, Anne, S.; Chiron Corporation, Intellectual Property - R440, P.O. Box 8097, Emeryville, CA 94662-8097 (US) et al.			
(54) Title: MODIFIED HIV ENV POLYPEPTIDES			
(57) Abstract			
<p>Polynucleotide encoding modified HIV Env polypeptides are disclosed. The Env polypeptides are modified so as to expose at least part of the CD4 binding region. Methods of diagnosis, treatment and prevention using the polynucleotides and polypeptides are also provided.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

MODIFIED HIV ENV POLYPEPTIDESTechnical Field

5       The invention relates generally to modified HIV envelope (Env) polypeptides which are useful as immunizing agents or for generating an immune response in a subject, for example a cellular immune response or a protective immune response. More particularly, the invention relates Env polypeptides such as gp120, gp140 or gp160, wherein at least one of the native  $\beta$ -sheet configurations has been modified. The invention also pertains to methods  
10 of using these polypeptides to elicit an immune response against a broad range of HIV subtypes.

Background of the Invention

      The human immunodeficiency virus (HIV-1, also referred to as HTLV-III, LAV or  
15 HTLV-III/LAV) is the etiological agent of the acquired immune deficiency syndrome (AIDS) and related disorders. (see, e.g., Barre-Sinoussi, et al., (1983) *Science* 220:868-871; Gallo et al. (1984) *Science* 224:500-503; Levy et al., (1984) *Science* 225:840-842; Siegal et al., (1981) *N. Engl. J. Med.* 305:1439-1444). AIDS patients usually have a long asymptomatic period followed by the progressive degeneration of the immune system and the central nervous  
20 system. Replication of the virus is highly regulated, and both latent and lytic infection of the CD4 positive helper subset of T-lymphocytes occur in tissue culture (Zagury et al., (1986) *Science* 231:850-853). Molecular studies of HIV-1 show that it encodes a number of genes (Ratner et al., (1985) *Nature* 313:277-284; Sanchez-Pescador et al., (1985) *Science* 227:484-492), including three structural genes -- gag, pol and env -- that are common to all  
25 retroviruses. Nucleotide sequences from viral genomes of other retroviruses, particularly HIV-2 and simian immunodeficiency viruses, SIV (previously referred to as STLV-III), also contain these structural genes. (Guyader et al., (1987) *Nature* 326:662-669; Chakrabarti et al., (1987) *Nature*

      The envelope protein of HIV-1, HIV-2 and SIV is a glycoprotein of about 160 kd  
30 (gp160). During virus infection of the host cell, gp160 is cleaved by host cell proteases to form gp120 and the integral membrane protein, gp41. The gp41 portion is anchored in the

membrane bilayer of virion, while the gp120 segment protrudes into the surrounding environment. gp120 and gp41 are more covalently associated and free gp120 can be released from the surface of virions and infected cells.

As depicted in Figure 1, crystallography studies of the gp120 core polypeptide indicate that this polypeptide is folded into two major domains having certain emanating structures. The inner domain (inner with respect to the N and C terminus) features a two-helix, two-stranded bundle with a small five-stranded  $\beta$ -sandwich at its termini-proximal end and a projection at the distal end from which the V1/V2 stem emanates. The outer domain is a staked double barrel that lies along side the inner domain so that the outer barrel and inner bundle axes are approximately parallel. Between the distal inner domain and the distal outer domain is a four-stranded bridging sheet which holds a peculiar minidomain in contact with, but distinct from, the inner, the outer domain, and the V1/V2 domain. The bridging sheet is composed of four  $\beta$ -strand structures ( $\beta$ -3,  $\beta$ -2,  $\beta$ -21,  $\beta$ -20, shown in Figure 1). The bridging region can be seen in Figure 1 packing primarily over the inner domain, although some surface residues of the outer domain, such as Phe 382, reach into the bridging sheet to form part of its hydrophobic core.

The basic unit of the  $\beta$ -sheet conformation of the bridging sheet region is the  $\beta$ -strand which exists as a less tightly coiled helix, with 2.0 residues per turn. The  $\beta$ -strand conformation is only stable when incorporated into a  $\beta$ -sheet, where hydrogen bonds with close to optimal geometry are formed between the peptide groups on adjacent  $\beta$ -strands; the dipole moments of the strands are also aligned favorably. Side chains from adjacent residues of the same strand protrude from opposite sides of the sheet and do not interact with each other, but have significant interactions with their backbone and with the side chains of neighboring strands. For a general description of  $\beta$ -sheets, see, e.g., T.E. Creighton, Proteins: Structures and Molecular Properties (W.H. Freeman and Company, 1993); and A.L. Lehninger, Biochemistry (Worth Publishers, Inc., 1975).

The gp120 polypeptide is instrumental in mediating entry into the host cell. Recent studies have indicated that binding of CD4 to gp120 induces a conformational change in Env that allows for binding to a co-receptor (e.g. a chemokine receptor) and subsequent entry of the virus into the cell. (Wyatt, R., et al. (1998) *Nature* 393:705-711; Kwong, P., et al. (1998) *Nature* 393:648-659). Referring again to Figure 1, CD4 is bound into a depression formed at the interface of the outer domain, the inner domain and the bridging sheet of gp120.

Immunogenicity of the gp120 polypeptide has also been studied. For example, individuals infected by HIV-1 usually develop antibodies that can neutralize the virus in *in vitro* assays, and this response is directed primarily against linear neutralizing determinants in the third variable loop of gp120 glycoprotein (Javaherian, K., et al. (1989) *Proc. Natl. Acad. Sci.* **86**:6786-6772; Matsushita, M., et al. (1988) *J. Virol.* **62**:2107-2144; Putney, S., et al. (1986) *Science* **234**:1392-1395; Rushe, J. R., et al. (1988) *Proc. Nat. Acad. Sci. USA* **85**:3198-3202.). However, these antibodies generally exhibit the ability to neutralize only a limited number of HIV-1 strains (Matthews, T. (1986) *Proc. Natl. Acad. Sci. USA* **83**:9709-9713; Nara, P. L., et al. (1988) *J. Virol.* **62**:2622-2628; Palker, T. J., et al. (1988) *Proc. Natl. Acad. Sci. USA* **85**:1932-1936). Later in the course of HIV infection in humans, antibodies capable of neutralizing a wider range of HIV-1 isolates appear (Barre-Sinoussi, F., et al. (1983) *Science* **220**:868-871; Robert-Guroff, M., et al. (1985) *Nature* (London) **316**:72-74; Weis, R., et al. (1985) *Nature* (London) **316**:69-72; Weis, R., et al. (1986) *Nature* (London) **324**:572-575).

Recent work done by Stamatatos et al (1998) *AIDS Res Hum Retroviruses* **14**(13):1129-39, shows that a deletion of the variable region 2 from a HIV-1<sub>SP162</sub> virus, which utilizes the CCR-5 co-receptor for virus entry, rendered the virus highly susceptible to serum-mediated neutralization. This V2 deleted virus was also neutralized by sera obtained from patients infected not only with clade B HIV-1 isolates but also with clade A, C, D and F HIV-1 isolates. However, deletion of the variable region 1 had no effect. Deletion of the variable regions 1 and 2 from a LAI isolate HIV-1<sub>MB</sub> also increased the susceptibility to neutralization by monoclonal antibodies whose epitopes are located within the V3 loop, the CD4-binding site, and conserved gp120 regions (Wyatt, R., et al. (1995) *J Virol.* **69**:5723-5733). Rabbit immunogenicity studies done with the HIV-1 virus with deletions in the V1/V2 and V3 region from the LAI strain, which uses the CXCR4 co-receptor for virus entry, showed no improvement in the ability of Env to raise neutralizing antibodies (Leu et al. (1998) *AIDS Res. and Human Retroviruses*. **14**:151-155).

Further, a subset of the broadly reactive antibodies, found in most infected individuals, interferes with the binding of gp120 and CD4 (Kang, C.-Y., et al. (1991) *Proc. Natl. Acad. Sci. USA* **88**:6171-6175; McDougal, J. S., et al. (1986) *J. Immunol.* **137**:2937-2944). Other antibodies are believed to bind to the chemokine receptor binding region after CD4 has bound to Env (Thali et al. (1993) *J. Virol.* **67**:3978-3988). The fact that neutralizing

antibodies generated during the course of HIV infection do not provide permanent antiviral effect may in part be due to the generation of "neutralization escapes" virus mutants and to the general decline in the host immune system associated with pathogenesis. In contrast, the presence of pre-existing neutralizing antibodies upon initial HIV-1 exposure will likely have a protective effect.

It is widely thought that a successful vaccine should be able to induce a strong, broadly neutralizing antibody response against diverse HIV-1 strains (Montefiori and Evans (1999) *AIDS Res. Hum. Ret.* 15(8):689-698; Bolognesi, D., P., et al. (1994) *Ann. Int. Med.* 8:603-611; Haynes, B., F., et al. (1996) *Science* ;271: 324-328.). Neutralizing antibodies, by attaching to the incoming virions, can reduce or even prevent their infectivity for target cells and prevent the cell-to-cell spread of virus in tissue culture (Hu et al. (1992) *Science* 255:456-459; Burton, D., R. and Montefiori, D. (1997) *AIDS* 11(suppl. A): 587-598). However as described above, antibodies directed against gp120 do not generally exhibit broad antibody responses against different HIV strains.

Currently, the focus of vaccine development, from the perspective of humoral immunity, is on the neutralization of primary isolates that utilize the CCR5 chemokine co-receptor believed to be important in virus entry (Zhu, T., et al. (1993) *Science* 261:1179-1181; Fiore, J., et al. (1994) *Virology*; 204:297-303). These viruses are generally much more resistant to antibody neutralization than T-cell line adapted strains that use the CXCR4 co-receptor, although both can be neutralized *in vitro* by certain broadly and potent acting monoclonal antibodies, such as IgG1b12, 2G12 and 2F5 (Trkola, A., et al. (1995) *J. Virol.* 69:6609-6617; D'Sousa PM., et al (1997) *J. Infect. Dis.* 175:1062-1075). These monoclonal antibodies are directed to the CD4 binding site, a glycosylation site and to the gp41 fusion domain, respectively. The problem that remains, however, is that it is not known how to induce antibodies of the appropriate specificity by vaccination. Antibodies (Abs) elicited by gp120 glycoprotein from a given isolate are usually only able to neutralize closely related viruses generally from similar, usually from the same, HIV-1 subtype.

Despite the above approaches, there remains a need for Env antigens that can elicit an immunological response (e.g., neutralizing and/or protective antibodies) in a subject against multiple HIV strains and subtypes, for example when administered as a vaccine. The present invention solves these and other problems by providing modified Env polypeptides (e.g., gp120) to expose epitopes in or near the CD4 binding site.

Summary of the Invention

In accordance with the present invention, modified HIV Env polypeptides are provided. In particular, deletions and/or mutations are made in one or more of the 4- $\beta$  antiparallel-bridging sheet in the HIV Env polypeptide. In this way, enough structure is left to allow correct folding of the polypeptide, for example of gp120, yet enough of the bridging sheet is removed to expose the CD4 groove, allowing an immune response to be generated against epitopes in or near the CD4 binding site of the Env polypeptide (*e.g.*, gp120).

In one aspect, the invention includes a polynucleotide encoding a modified HIV Env polypeptide wherein the polypeptide has at least one modified (*e.g.*, deleted or replaced) amino acid residue deleted in the region corresponding to residues 421 to 436 relative to HXB-2, for example the constructs depicted in Figures 6-29 (SEQ ID NOs:3 to 26). In certain embodiments, the polynucleotide also has the region corresponding to residues 124-198 of the polypeptide HXB-2 (*e.g.*, V1/V2) deleted and at least one amino acid deleted or replaced in the regions corresponding to the residues 119 to 123 and 199 to 210, relative to HXB-2. In other embodiments, these polynucleotides encode Env polypeptides having at least one amino acid of the small loop of the bridging sheet (*e.g.*, amino acid residues 427 to 429 relative to HXB-2) deleted or replaced. The amino acid sequences of the modified polypeptides encoded by the polynucleotides of the present invention can be based on any HIV variant, for example SF162.

In another aspect, the invention includes immunogenic modified HIV Env polypeptides having at least one modified (*e.g.*, deleted or replaced) amino acid residue deleted in the region corresponding to residues 421 to 436 relative to HXB-2, for example a deletion or replacement of one amino acids in the small loop region (*e.g.*, amino acid residues 427 to 429 relative to HXB-2). These polypeptides may have modifications (*e.g.*, a deletion or a replacement) of at least one amino acid between about amino acid residue 420 and amino acid residue 436, relative to HXB-2 and, optionally, may have deletions or truncations of the V1 and/or V2 regions. The immunogenic, modified polypeptides of the present invention can be based on any HIV variant, for example SF162.

In another aspect, the invention includes a vaccine composition comprising any of the polynucleotides encoding modified Env polypeptides described above. Vaccine compositions comprising the modified Env polypeptides and, optionally, an adjuvant are also included in the invention.

In yet another aspect, the invention includes a method of inducing an immune response in subject comprising, administering one or more of the polynucleotides or constructs described above in an amount sufficient to induce an immune response in the subject. In certain embodiments, the method further comprises administering an adjuvant to the subject.

In another aspect, the invention includes a method of inducing an immune response in a subject comprising administering a composition comprising any of the modified Env polypeptides described above and an adjuvant. The composition is administered in an amount sufficient to induce an immune response in the subject.

In another aspect, the invention includes a method of inducing an immune response in a subject comprising

(a) administering a first composition comprising any of the polynucleotides described above in a priming step and

(b) administering a second composition comprising any of the modified Env polypeptides described above, as a booster, in an amount sufficient to induce an immune response in the subject. In certain embodiments, the first composition, the second composition or both the first and second compositions further comprise an adjuvant.

These and other embodiments of the subject invention will readily occur to those of skill in the art in light of the disclosure herein.

#### Brief Description of the Drawings

Figure 1 is a schematic depiction of the tertiary structure of the HIV-1<sub>HXB-2</sub> Env gp120 polypeptide, as determined by crystallography studies.

Figures 2A-C depict alignment of the amino acid sequence of wild-type HIV-1<sub>HXB-2</sub> Env gp160 polypeptide (SEQ ID NO:1) with amino acid sequence of HIV variants SF162 (shown as "162") (SEQ ID NO:2), SF2, CM236 and US4. Arrows indicate the regions that are deleted or replaced in the modified polypeptides. Black dots indicate conserved cysteine residues. The star indicates the position of the last amino acid in gp120.

Figures 3A-J depict alignment of nucleotide sequences of polynucleotides encoding modified Env polypeptides having V1/V2 deletions. The unmodified amino acid residues encoded by these sequences correspond to wildtype SF162 residues but are numbered relative to HXB-2.



Figures 4A-M depict alignment of nucleotide sequences of polynucleotides encoding modified Env polypeptides having deletions or replacements in the small loop. The unmodified amino acid residues encoded by these sequences correspond to wildtype SF162 residues but are numbered relative to HXB-2.

5        Figures 5A-N depict alignment of nucleotide sequences of polynucleotides encoding modified Env polypeptides having both V1/V2 deletions and, in addition, deletions or replacements in the small loop. The unmodified amino acid residues encoded by these sequences correspond to wildtype SF162 residues but are numbered relative to HXB-2.

10        Figure 6 depicts the nucleotide sequence of the construct designated Val120-Ala204 (SEQ ID NO:3).

      Figure 7 depicts the nucleotide sequence of the construct designated Val120-Ile201 (SEQ ID NO:4).

      Figure 8 depicts the nucleotide sequence of the construct designated Val120-Ile201B (SEQ ID NO:5).

15        Figure 9 depicts the nucleotide sequence of the construct designated Lys121-Val200 (SEQ ID NO:6).

      Figure 10 depicts the nucleotide sequence of the construct designated Leu122-Ser199 (SEQ ID NO:7).

20        Figure 11 depicts the nucleotide sequence of the construct designated Val120-Thr202 (SEQ ID NO:8).

      Figure 12 depicts the nucleotide sequence of the construct designated Trp427-Gly431 (SEQ ID NO:9).

      Figure 13 depicts the nucleotide sequence of the construct designated Arg426-Gly431 (SEQ ID NO:10).

25        Figure 14 depicts the nucleotide sequence of the construct designated Arg426-Gly431B (SEQ ID NO:11).

      Figure 15 depicts the nucleotide sequence of the construct designated Arg426-Lys432 (SEQ ID NO:12).

30        Figure 16 depicts the nucleotide sequence of the construct designated Asn425-Lys432 (SEQ ID NO:13).

      Figure 17 depicts the nucleotide sequence of the construct designated Ile424-Ala433 (SEQ ID NO:14).

Figure 18 depicts the nucleotide sequence of the construct designated Ile423-Met434 (SEQ ID NO:15).

Figure 19 depicts the nucleotide sequence of the construct designated Gln422-Tyr435 (SEQ ID NO:16).

5        Figure 20 depicts the nucleotide sequence of the construct designated Gln422-Tyr435B (SEQ ID NO:17).

Figure 21 depicts the nucleotide sequence of the construct designated Leu122-Ser199;Arg426-Gly431 (SEQ ID NO:18).

10       Figure 22 depicts the nucleotide sequence of the construct designated Leu122-Ser199;Arg426-Lys432 (SEQ ID NO:19).

Figure 23 depicts the nucleotide sequence of the construct designated Leu122-Ser199;Trp427-Gly431 (SEQ ID NO:20).

Figure 24 depicts the nucleotide sequence of the construct designated Lys121-Val200;Asn425-Lys432 (SEQ ID NO:21).

15       Figure 25 depicts the nucleotide sequence of the construct designated Val120-Ile201;Ile424-Ala433 (SEQ ID NO:22).

Figure 26 depicts the nucleotide sequence of the construct designated Val120-Ile201B; Ile424-Ala433 (SEQ ID NO:23).

20       Figure 27 depicts the nucleotide sequence of the construct designated Val120-Thr202;Ile424-Ala433 (SEQ ID NO:24).

Figure 28 depicts the nucleotide sequence of the construct designated Val127-Asn195 (SEQ ID NO:25).

25       Figure 29 depicts the nucleotide sequence of the construct designated Val127-Asn195; Arg426-Gly431 (SEQ ID NO:26).

#### Detailed Description of the Invention

The practice of the present invention will employ, unless otherwise indicated, conventional methods of protein chemistry, viral immunobiology, molecular biology and recombinant DNA techniques within the skill of the art. Such techniques are explained fully  
30       in the literature. See, e.g., T.E. Creighton, Proteins: Structures and Molecular Properties (W.H. Freeman and Company, 1993); Nelson L.M. and Jerome H.K. HIV Protocols in Methods in Molecular Medicine, vol. 17, 1999; Sambrook, et al., Molecular Cloning: A

Laboratory Manual (Cold Spring Harbor Laboratory, 1989); F.M. Ausubel et al. Current Protocols in Molecular Biology, Greene Publishing Associates & Wiley Interscience New York; and Lipkowitz and Boyd, Reviews in Computational Chemistry, volumes 1-present (Wiley-VCH, New York, New York, 1999).

5 It must be noted that, as used in this specification and the appended claims, the singular forms "a", "an" and "the" include plural referents unless the content clearly dictates otherwise. Thus, for example, reference to "a polypeptide" includes a mixture of two or more polypeptides, and the like.

## 10 Definitions

In describing the present invention, the following terms will be employed, and are intended to be defined as indicated below.

The terms "polypeptide," and "protein" are used interchangeably herein to denote any polymer of amino acid residues. The terms encompass peptides, oligopeptides, dimers, 15 multimers, and the like. Such polypeptides can be derived from natural sources or can be synthesized or recombinantly produced. The terms also include postexpression modifications of the polypeptide, for example, glycosylation, acetylation, phosphorylation, etc.

A polypeptide as defined herein is generally made up of the 20 natural amino acids Ala (A), Arg (R), Asn (N), Asp (D), Cys (C), Gln (Q), Glu (E), Gly (G), His (H), Ile (I), Leu 20 (L), Lys (K), Met (M), Phe (F), Pro (P), Ser (S), Thr (T), Trp (W), Tyr (Y) and Val (V) and may also include any of the several known amino acid analogs, both naturally occurring and synthesized analogs, such as but not limited to homoisoleucine, asaleucine, 2-(methylenecyclopropyl)glycine, S-methylcysteine, S-(prop-1-enyl)cysteine, homoserine, ornithine, norleucine, norvaline, homoarginine, 3-(3-carboxyphenyl)alanine, 25 cyclohexylalanine, mimosine, pipecolic acid, 4-methylglutamic acid, canavanine, 2,3-diaminopropionic acid, and the like. Further examples of polypeptide agents which will find use in the present invention are set forth below.

By "geometry" or "tertiary structure" of a polypeptide or protein is meant the overall 3-D configuration of the protein. As described herein, the geometry can be determined, for 30 example, by crystallography studies or by using various programs or algorithms which predict the geometry based on interactions between the amino acids making up the primary and secondary structures.

By "wild type" polypeptide, polypeptide agent or polypeptide drug, is meant a naturally occurring polypeptide sequence, and its corresponding secondary structure. An "isolated" or "purified" protein or polypeptide is a protein which is separate and discrete from a whole organism with which the protein is normally associated in nature. It is apparent that the term denotes proteins of various levels of purity. Typically, a composition containing a purified protein will be one in which at least about 35%, preferably at least about 40-50%, more preferably, at least about 75-85%, and most preferably at least about 90% or more, of the total protein in the composition will be the protein in question.

By "Env polypeptide" is meant a molecule derived from an envelope protein, preferably from HIV Env. The envelope protein of HIV-1 is a glycoprotein of about 160 kd (gp160). During virus infection of the host cell, gp160 is cleaved by host cell proteases to form gp120 and the integral membrane protein, gp41. The gp41 portion is anchored in (and spans) the membrane bilayer of virion, while the gp120 segment protrudes into the surrounding environment. As there is no covalent attachment between gp120 and gp41, free gp120 is released from the surface of virions and infected cells. Env polypeptides may also include gp140 polypeptides. Env polypeptides can exist as monomers, dimers or multimers.

By a "gp120 polypeptide" is meant a molecule derived from a gp120 region of the Env polypeptide. Preferably, the gp120 polypeptide is derived from HIV Env. The primary amino acid sequence of gp120 is approximately 511 amino acids, with a polypeptide core of about 60,000 daltons. The polypeptide is extensively modified by N-linked glycosylation to increase the apparent molecular weight of the molecule to 120,000 daltons. The amino acid sequence of gp120 contains five relatively conserved domains interspersed with five hypervariable domains. The positions of the 18 cysteine residues in the gp120 primary sequence of the HIV-1<sub>HXB-2</sub> (hereinafter "HXB-2") strain, and the positions of 13 of the approximately 24 N-linked glycosylation sites in the gp120 sequence are common to most, if not all, gp120 sequences. The hypervariable domains contain extensive amino acid substitutions, insertions and deletions. Despite this variation, most, if not all, gp120 sequences preserve the virus's ability to bind to the viral receptor CD4. A "gp120 polypeptide" includes both single subunits or multimers.

Env polypeptides (e.g., gp120, gp140 and gp160) include a "bridging sheet" comprised of 4 anti-parallel  $\beta$ -strands ( $\beta$ -2,  $\beta$ -3,  $\beta$ -20 and  $\beta$ -21) that form a  $\beta$ -sheet. Extruding from one pair of the  $\beta$ -strands ( $\beta$ -2 and  $\beta$ -3) are two loops, V1 and V2. The  $\beta$ -2

sheet occurs at approximately amino acid residue 119 (Cys) to amino acid residue 123 (Thr) while  $\beta$ -3 occurs at approximately amino acid residue 199 (Ser) to amino acid residue 201 (Ile), relative to HXB-2. The "V1/V2 region" occurs at approximately amino acid positions 126 (Cys) to residue 196 (Cys), relative to HXB-2. (see, e.g., Wyatt et al. (1995) *J. Virol.* 69:5723-5733; Stamatas et al. (1998) *J. Virol.* 72:7840-7845). Extruding from the second pair of  $\beta$ -strands ( $\beta$ -20 and  $\beta$ -21) is a "small-loop" structure, also referred to herein as "the bridging sheet small loop." In HXB-2,  $\beta$ -20 extends from about amino acid residue 422 (Gln) to amino acid residue 426 (Met) while  $\beta$ -21 extends from about amino acid residue 430 (Val) to amino acid residue 435 (Tyr). In variant SF162, the Met-426 is an Arg (R) residue.

The "small loop" extends from about amino acid residue 427 (Trp) through 429 (Lys), relative to HXB-2. A representative diagram of gp120 showing the bridging sheet, the small loop, and V1/V2 is shown in Figure 1. In addition, alignment of the amino acid sequences of Env polypeptide gp160 of selected variants is shown, relative to HXB-2, in Figures 2A-C.

Furthermore, an "Env polypeptide" or "gp120 polypeptide" as defined herein is not limited to a polypeptide having the exact sequence described herein. Indeed, the HIV genome is in a state of constant flux and contains several variable domains which exhibit relatively high degrees of variability between isolates. It is readily apparent that the terms encompass Env (e.g., gp120) polypeptides from any of the identified HIV isolates, as well as newly identified isolates, and subtypes of these isolates. Descriptions of structural features are given herein with reference to HXB-2. One of ordinary skill in the art in view of the teachings of the present disclosure and the art can determine corresponding regions in other HIV variants (e.g., isolates HIV<sub>IIIb</sub>, HIV<sub>SF2</sub>, HIV-1<sub>SF162</sub>, HIV-1<sub>SF170</sub>, HIV<sub>LA</sub>, HIV<sub>LA1</sub>, HIV<sub>MN</sub>, HIV-1<sub>CM235</sub>, HIV-1<sub>US4</sub>, other HIV-1 strains from diverse subtypes (e.g., subtypes, A through G, and O), HIV-2 strains and diverse subtypes (e.g., HIV-2<sub>UC1</sub> and HIV-2<sub>UC2</sub>), and simian immunodeficiency virus (SIV). (See, e.g., Virology, 3rd Edition (W.K. Joklik ed. 1988); *Fundamental Virology*, 2nd Edition (B.N. Fields and D.M. Knipe, eds. 1991); *Virology*, 3rd Edition (Fields, BN, DM Knipe, PM Howley, Editors, 1996, Lippincott-Raven, Philadelphia, PA; for a description of these and other related viruses), using for example, sequence comparison programs (e.g., BLAST and others described herein) or identification and alignment of structural features (e.g., a program such as the "ALB" program described herein that can identify  $\beta$ -sheet regions). The actual amino acid sequences of the modified Env polypeptides can be based on any HIV variant.

Additionally, the term "Env polypeptide" (*e.g.*, "gp120 polypeptide") encompasses proteins which include additional modifications to the native sequence, such as additional internal deletions, additions and substitutions. These modifications may be deliberate, as through site-directed mutagenesis, or may be accidental, such as through naturally occurring mutational events. Thus, for example, if the Env polypeptide is to be used in vaccine compositions, the modifications must be such that immunological activity (*i.e.*, the ability to elicit an antibody response to the polypeptide) is not lost. Similarly, if the polypeptides are to be used for diagnostic purposes, such capability must be retained.

Thus, a "modified Env polypeptide" is an Env polypeptide (*e.g.*, gp120 as defined above), which has been manipulated to delete or replace all or a part of the bridging sheet portion and, optionally, the variable regions V1 and V2. Generally, modified Env (*e.g.*, gp120) polypeptides have enough of the bridging sheet removed to expose the CD4 binding site, but leave enough of the structure to allow correct folding (*e.g.*, correct geometry). Thus, modifications to the  $\beta$ -20 and  $\beta$ -21 regions (between about amino acid residues 420 and 435 relative to HXB-2) are preferred. Additionally, modifications to the  $\beta$ -2 and  $\beta$ -3 regions (between about amino acid residues 119 (Cys) and 201 (Ile)) and modifications (*e.g.*, truncations) to the V1 and V2 loop regions may also be made. Although not all possible  $\beta$ -sheet and V1/V2 modifications have been exemplified herein, it is to be understood that other disrupting modifications are also encompassed by the present invention.

Normally, such a modified polypeptide is capable of secretion into growth medium in which an organism expressing the protein is cultured. However, for purposes of the present invention, such polypeptides may also be recovered intracellularly. Secretion into growth media is readily determined using a number of detection techniques, including, *e.g.*, polyacrylamide gel electrophoresis and the like, and immunological techniques such as Western blotting and immunoprecipitation assays as described in, *e.g.*, International Publication No. WO 96/04301, published February 15, 1996.

A gp120 or other Env polypeptide is produced "intracellularly" when it is found within the cell, either associated with components of the cell, such as in association with the endoplasmic reticulum (ER) or the Golgi Apparatus, or when it is present in the soluble cellular fraction. The gp120 and other Env polypeptides of the present invention may also be secreted into growth medium so long as sufficient amounts of the polypeptides remain

present within the cell such that they can be purified from cell lysates using techniques described herein.

An "immunogenic" gp120 or other Env protein is a molecule that includes at least one epitope such that the molecule is capable of either eliciting an immunological reaction in an individual to which the protein is administered or, in the diagnostic context, is capable of reacting with antibodies directed against the HIV in question.

By "epitope" is meant a site on an antigen to which specific B cells and/or T cells respond, rendering the molecule including such an epitope capable of eliciting an immunological reaction or capable of reacting with HIV antibodies present in a biological sample. The term is also used interchangeably with "antigenic determinant" or "antigenic determinant site." An epitope can comprise 3 or more amino acids in a spatial conformation unique to the epitope. Generally, an epitope consists of at least 5 such amino acids and, more usually, consists of at least 8-10 such amino acids. Methods of determining spatial conformation of amino acids are known in the art and include, for example, x-ray crystallography and 2-dimensional nuclear magnetic resonance. Furthermore, the identification of epitopes in a given protein is readily accomplished using techniques well known in the art, such as by the use of hydrophobicity studies and by site-directed serology. See, also, Geysen et al., *Proc. Natl. Acad. Sci. USA* (1984) 81:3998-4002 (general method of rapidly synthesizing peptides to determine the location of immunogenic epitopes in a given antigen); U.S. Patent No. 4,708,871 (procedures for identifying and chemically synthesizing epitopes of antigens); and Geysen et al., *Molecular Immunology* (1986) 23:709-715 (technique for identifying peptides with high affinity for a given antibody). Antibodies that recognize the same epitope can be identified in a simple immunoassay showing the ability of one antibody to block the binding of another antibody to a target antigen.

An "immunological response" or "immune response" as used herein is the development in the subject of a humoral and/or a cellular immune response to the Env (e.g., gp120) polypeptide when the polypeptide is present in a vaccine composition. These antibodies may also neutralize infectivity, and/or mediate antibody-complement or antibody dependent cell cytotoxicity to provide protection to an immunized host. Immunological reactivity may be determined in standard immunoassays, such as a competition assays, well known in the art.

Techniques for determining amino acid sequence "similarity" are well known in the art. In general, "similarity" means the exact amino acid to amino acid comparison of two or more polypeptides at the appropriate place, where amino acids are identical or possess similar chemical and/or physical properties such as charge or hydrophobicity. A so-termed "percent similarity" then can be determined between the compared polypeptide sequences.

Techniques for determining nucleic acid and amino acid sequence identity also are well known in the art and include determining the nucleotide sequence of the mRNA for that gene (usually via a cDNA intermediate) and determining the amino acid sequence encoded thereby, and comparing this to a second amino acid sequence. In general, "identity" refers to an exact nucleotide to nucleotide or amino acid to amino acid correspondence of two polynucleotides or polypeptide sequences, respectively.

Two or more polynucleotide sequences can be compared by determining their "percent identity." Two or more amino acid sequences likewise can be compared by determining their "percent identity." The percent identity of two sequences, whether nucleic acid or peptide sequences, is generally described as the number of exact matches between two aligned sequences divided by the length of the shorter sequence and multiplied by 100. An approximate alignment for nucleic acid sequences is provided by the local homology algorithm of Smith and Waterman, *Advances in Applied Mathematics* 2:482-489 (1981). This algorithm can be extended to use with peptide sequences using the scoring matrix developed by Dayhoff, *Atlas of Protein Sequences and Structure*, M.O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA, and normalized by Gribskov, *Nucl. Acids Res.* 14(6):6745-6763 (1986). An implementation of this algorithm for nucleic acid and peptide sequences is provided by the Genetics Computer Group (Madison, WI) in their BestFit utility application. The default parameters for this method are described in the Wisconsin Sequence Analysis Package Program Manual, Version 8 (1995) (available from Genetics Computer Group, Madison, WI). Other equally suitable programs for calculating the percent identity or similarity between sequences are generally known in the art.

For example, percent identity of a particular nucleotide sequence to a reference sequence can be determined using the homology algorithm of Smith and Waterman with a default scoring table and a gap penalty of six nucleotide positions. Another method of establishing percent identity in the context of the present invention is to use the MPSRCH



package of programs copyrighted by the University of Edinburgh, developed by John F. Collins and Shane S. Sturrok, and distributed by IntelliGenetics, Inc. (Mountain View, CA). From this suite of packages, the Smith-Waterman algorithm can be employed where default parameters are used for the scoring table (for example, gap open penalty of 12, gap extension penalty of one, and a gap of six). From the data generated, the "Match" value reflects "sequence identity." Other suitable programs for calculating the percent identity or similarity between sequences are generally known in the art, such as the alignment program BLAST, which can also be used with default parameters. For example, BLASTN and BLASTP can be used with the following default parameters: genetic code = standard; filter = none; strand = both; cutoff = 60; expect = 10; Matrix = BLOSUM62; Descriptions = 50 sequences; sort by = HIGH SCORE; Databases = non-redundant, GenBank + EMBL + DDBJ + PDB + GenBank CDS translations + Swiss protein + Spupdate + PIR. Details of these programs can be found at the following internet address: <http://www.ncbi.nlm.gov/cgi-bin/BLAST>.

One of skill in the art can readily determine the proper search parameters to use for a given sequence in the above programs. For example, the search parameters may vary based on the size of the sequence in question. Thus, for example, a representative embodiment of the present invention would include an isolated polynucleotide having X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least about 50% identity to Y contiguous nucleotides derived from any of the sequences described herein, (ii) X equals Y, and (iii) X is greater than or equal to 6 nucleotides and up to 5000 nucleotides, preferably greater than or equal to 8 nucleotides and up to 5000 nucleotides, more preferably 10-12 nucleotides and up to 5000 nucleotides, and even more preferably 15-20 nucleotides, up to the number of nucleotides present in the full-length sequences described herein (e.g., see the Sequence Listing and claims), including all integer values falling within the above-described ranges.

The synthetic expression cassettes (and purified polynucleotides) of the present invention include related polynucleotide sequences having about 80% to 100%, greater than 80-85%, preferably greater than 90-92%, more preferably greater than 95%, and most preferably greater than 98% sequence (including all integer values falling within these described ranges) identity to the synthetic expression cassette sequences disclosed herein (for example, to the claimed sequences or other sequences of the present invention) when the sequences of the present invention are used as the query sequence.

Computer programs are also available to determine the likelihood of certain polypeptides to form structures such as  $\beta$ -sheets. One such program, described herein, is the "ALB" program for protein and polypeptide secondary structure calculation and predication. In addition, secondary protein structure can be predicted from the primary amino acid sequence, for example using protein crystal structure and aligning the protein sequence related to the crystal structure (e.g., using Molecular Operating Environment (MOE) programs available from the Chemical Computing Group Inc., Montreal, P.Q., Canada). Other methods of predicting secondary structures are described, for example, in Garnier et al. (1996) *Methods Enzymol.* 266:540-553; Geourjon et al. (1995) *Comput. Applic. Biosci.* 11:681-684; Levin (1997) *Protein Eng.* 10:771-776; and Rost et al. (1993) *J. Molec. Biol.* 232:584-599.

Homology can also be determined by hybridization of polynucleotides under conditions which form stable duplexes between homologous regions, followed by digestion with single-stranded-specific nuclease(s), and size determination of the digested fragments. Two DNA, or two polypeptide sequences are "substantially homologous" to each other when the sequences exhibit at least about 80%-85%, preferably at least about 90%, and most preferably at least about 95%-98% sequence identity over a defined length of the molecules, as determined using the methods above. As used herein, substantially homologous also refers to sequences showing complete identity to the specified DNA or polypeptide sequence. DNA sequences that are substantially homologous can be identified in a Southern hybridization experiment under, for example, stringent conditions, as defined for that particular system. Defining appropriate hybridization conditions is within the skill of the art. See, e.g., Sambrook et al., *supra*; *DNA Cloning, supra*; *Nucleic Acid Hybridization, supra*.

A "coding sequence" or a sequence which "encodes" a selected protein, is a nucleic acid sequence which is transcribed (in the case of DNA) and translated (in the case of mRNA) into a polypeptide *in vitro* or *in vivo* when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxy) terminus. A coding sequence can include, but is not limited to cDNA from viral nucleotide sequences as well as synthetic and semisynthetic DNA sequences and sequences including base analogs. A transcription termination sequence may be located 3' to the coding sequence.

"Control elements" refers collectively to promoter sequences, ribosome binding sites, polyadenylation signals, transcription termination sequences, upstream regulatory domains, enhancers, and the like, which collectively provide for the transcription and translation of a coding sequence in a host cell. Not all of these control elements need always be present so long as the desired gene is capable of being transcribed and translated.

A control element "directs the transcription" of a coding sequence in a cell when RNA polymerase will bind the promoter sequence and transcribe the coding sequence into mRNA, which is then translated into the polypeptide encoded by the coding sequence.

"Operably linked" refers to an arrangement of elements wherein the components so described are configured so as to perform their usual function. Thus, control elements operably linked to a coding sequence are capable of effecting the expression of the coding sequence when RNA polymerase is present. The control elements need not be contiguous with the coding sequence, so long as they function to direct the expression thereof. Thus, for example, intervening untranslated yet transcribed sequences can be present between, e.g., a promoter sequence and the coding sequence and the promoter sequence can still be considered "operably linked" to the coding sequence.

"Recombinant" as used herein to describe a nucleic acid molecule means a polynucleotide of genomic, cDNA, semisynthetic, or synthetic origin which, by virtue of its origin or manipulation: (1) is not associated with all or a portion of the polynucleotide with which it is associated in nature; and/or (2) is linked to a polynucleotide other than that to which it is linked in nature. The term "recombinant" as used with respect to a protein or polypeptide means a polypeptide produced by expression of a recombinant polynucleotide. "Recombinant host cells," "host cells," "cells," "cell lines," "cell cultures," and other such terms denoting procaryotic microorganisms or eucaryotic cell lines cultured as unicellular entities, are used interchangeably, and refer to cells which can be, or have been, used as recipients for recombinant vectors or other transfer DNA, and include the progeny of the original cell which has been transfected. It is understood that the progeny of a single parental cell may not necessarily be completely identical in morphology or in genomic or total DNA complement to the original parent, due to accidental or deliberate mutation. Progeny of the parental cell which are sufficiently similar to the parent to be characterized by the relevant property, such as the presence of a nucleotide sequence encoding a desired peptide, are included in the progeny intended by this definition, and are covered by the above terms.

By "vertebrate subject" is meant any member of the subphylum chordata, including, without limitation, humans and other primates, including non-human primates such as chimpanzees and other apes and monkey species; farm animals such as cattle, sheep, pigs, goats and horses; domestic mammals such as dogs and cats; laboratory animals including rodents such as mice, rats and guinea pigs; birds, including domestic, wild and game birds such as chickens, turkeys and other gallinaceous birds, ducks, geese, and the like. The term does not denote a particular age. Thus, both adult and newborn individuals are intended to be covered.

As used herein, a "biological sample" refers to a sample of tissue or fluid isolated from an individual, including but not limited to, for example, blood, plasma, serum, fecal matter, urine, bone marrow, bile, spinal fluid, lymph fluid, samples of the skin, external secretions of the skin, respiratory, intestinal, and genitourinary tracts, samples derived from the gastric epithelium and gastric mucosa, tears, saliva, milk, blood cells, organs, biopsies and also samples of *in vitro* cell culture constituents including but not limited to conditioned media resulting from the growth of cells and tissues in culture medium, e.g., recombinant cells, and cell components.

The terms "label" and "detectable label" refer to a molecule capable of detection, including, but not limited to, radioactive isotopes, fluorescers, chemiluminescers, enzymes, enzyme substrates, enzyme cofactors, enzyme inhibitors, chromophores, dyes, metal ions, metal sols, ligands (e.g., biotin or haptens) and the like. The term "fluorescer" refers to a substance or a portion thereof which is capable of exhibiting fluorescence in the detectable range. Particular examples of labels which may be used with the invention include, but are not limited to fluorescein, rhodamine, dansyl, umbelliferone, Texas red, luminol, acridinium esters, NADPH,  $\alpha$ - $\beta$ -galactosidase, horseradish peroxidase, glucose oxidase, alkaline phosphatase and urease.

### Overview

The present invention concerns modified Env polypeptide molecules (e.g., glycoprotein ("gp") 120). Without being bound by a particular theory, it appears that it has been difficult to generate immunological responses against Env because the CD4 binding site is buried between the outer domain, the inner domain and the V1/V2 domains. Thus, although deletion of the V1/V2 domain may render the virus more susceptible to

neutralization by monoclonal antibody directed to the CD4 site, the bridging sheet covering most of the CD4 binding domain may prevent an antibody response. Thus, the present invention provides Env polypeptides that maintain their general overall structure yet expose the CD4 binding domain. This allows the generation of an immune response (*e.g.*, an antibody response) to epitopes in or near the CD4 binding site.

Various forms of the different embodiments of the invention, described herein, may be combined.

### **$\beta$ -Sheet Conformations**

In the present invention, location of the  $\beta$ -sheet structures were identified relative to 3-D (crystal) structure of an HXB-2 crystallized Env protein (see, Example 1A). Based on this structure, constructs encoding polypeptides having replacements and or excisions which maintain overall geometry while exposing the CD4 binding site were designed. In particular, the crystal structure of HXB-2 was downloaded from the Brookhaven Database. Using the default parameters of the Loop Search feature of the Biopolymer module of the Sybyl molecular modeling package, homology and fit of amino acids which could replace the native loops between  $\beta$ -strands yet maintain overall tertiary structure were determined. Constructs encoding the modified Env polypeptides were then designed (Example 1.B.).

Thus, the modified Env polypeptides typically have enough of the bridging sheet removed to expose the CD4 groove, but have enough of the structure to allow correct folding of the Env glycoprotein. Exemplary constructs are described below.

### **Polypeptide Production**

The polypeptides of the present invention can be produced in any number of ways which are well known in the art.

In one embodiment, the polypeptides are generated using recombinant techniques, well known in the art. In this regard, oligonucleotide probes can be devised based on the known sequences of the Env (*e.g.*, gp120) polypeptide genome and used to probe genomic or cDNA libraries for Env genes. The gene can then be further isolated using standard techniques and, *e.g.*, restriction enzymes employed to truncate the gene at desired portions of the full-length sequence. Similarly, the Env gene(s) can be isolated directly from cells and tissues containing the same, using known techniques, such as phenol extraction and the

sequence further manipulated to produce the desired truncations. See, e.g., Sambrook et al., *supra*, for a description of techniques used to obtain and isolate DNA.

The genes encoding the modified (e.g., truncated and/or substituted) polypeptides can be produced synthetically, based on the known sequences. The nucleotide sequence can be designed with the appropriate codons for the particular amino acid sequence desired. The complete sequence is generally assembled from overlapping oligonucleotides prepared by standard methods and assembled into a complete coding sequence. See, e.g., Edge (1981) *Nature* 292:756; Nambair et al. (1984) *Science* 223:1299; Jay et al. (1984) *J. Biol. Chem.* 259:6311; Stemmer et al. (1995) *Gene* 164:49-53.

Recombinant techniques are readily used to clone a gene encoding an Env polypeptide gene which can then be mutagenized *in vitro* by the replacement of the appropriate base pair(s) to result in the codon for the desired amino acid. Such a change can include as little as one base pair, effecting a change in a single amino acid, or can encompass several base pair changes. Alternatively, the mutations can be effected using a mismatched primer which hybridizes to the parent nucleotide sequence (generally cDNA corresponding to the RNA sequence), at a temperature below the melting temperature of the mismatched duplex. The primer can be made specific by keeping primer length and base composition within relatively narrow limits and by keeping the mutant base centrally located. See, e.g., Innis et al, (1990) *PCR Applications: Protocols for Functional Genomics*; Zoller and Smith, *Methods Enzymol.* (1983) 100:468. Primer extension is effected using DNA polymerase, the product cloned and clones containing the mutated DNA, derived by segregation of the primer extended strand, selected. Selection can be accomplished using the mutant primer as a hybridization probe. The technique is also applicable for generating multiple point mutations. See, e.g., Dalbie-McFarland et al. *Proc. Natl. Acad. Sci USA* (1982) 79:6409.

Once coding sequences for the desired proteins have been isolated or synthesized, they can be cloned into any suitable vector or replicon for expression. As will be apparent from the teachings herein, a wide variety of vectors encoding modified polypeptides can be generated by creating expression constructs which operably link, in various combinations, polynucleotides encoding Env polypeptides having deletions or mutation therein. Thus, polynucleotides encoding a particular deleted V1/V2 region can be operably linked with polynucleotides encoding polypeptides having deletions or replacements in the small loop

region and the construct introduced into a host cell for polypeptide expression. Non-limiting examples of such combinations are discussed in the Examples.

Numerous cloning vectors are known to those of skill in the art, and the selection of an appropriate cloning vector is a matter of choice. Examples of recombinant DNA vectors for cloning and host cells which they can transform include the bacteriophage  $\lambda$  (*E. coli*), pBR322 (*E. coli*), pACYC177 (*E. coli*), pKT230 (gram-negative bacteria), pGV1106 (gram-negative bacteria), pLAFR1 (gram-negative bacteria), pME290 (non-*E. coli* gram-negative bacteria), pHV14 (*E. coli* and *Bacillus subtilis*), pBD9 (*Bacillus*), pIJ61 (*Streptomyces*), pUC6 (*Streptomyces*), Ylp5 (*Saccharomyces*), YCp19 (*Saccharomyces*) and bovine papilloma virus (mammalian cells). See, generally, *DNA Cloning*: Vols. I & II, *supra*; Sambrook *et al.*, *supra*; B. Perbal, *supra*.

Insect cell expression systems, such as baculovirus systems, can also be used and are known to those of skill in the art and described in, e.g., Summers and Smith, *Texas Agricultural Experiment Station Bulletin No. 1555* (1987). Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *inter alia*, Invitrogen, San Diego CA ("MaxBac" kit).

Plant expression systems can also be used to produce the modified Env proteins. Generally, such systems use virus-based vectors to transfect plant cells with heterologous genes. For a description of such systems see, e.g., Porta *et al.*, *Mol. Biotech.* (1996) 5:209-221; and Hackland *et al.*, *Arch. Virol.* (1994) 139:1-22.

Viral systems, such as a vaccinia based infection/transfection system, as described in Tomei *et al.*, *J. Virol.* (1993) 67:4017-4026 and Selby *et al.*, *J. Gen. Virol.* (1993) 74:1103-1113, will also find use with the present invention. In this system, cells are first transfected *in vitro* with a vaccinia virus recombinant that encodes the bacteriophage T7 RNA polymerase. This polymerase displays exquisite specificity in that it only transcribes templates bearing T7 promoters. Following infection, cells are transfected with the DNA of interest, driven by a T7 promoter. The polymerase expressed in the cytoplasm from the vaccinia virus recombinant transcribes the transfected DNA into RNA which is then translated into protein by the host translational machinery. The method provides for high level, transient, cytoplasmic production of large quantities of RNA and its translation product(s).

The gene can be placed under the control of a promoter, ribosome binding site (for bacterial expression) and, optionally, an operator (collectively referred to herein as "control" elements), so that the DNA sequence encoding the desired Env polypeptide is transcribed into RNA in the host cell transformed by a vector containing this expression construction. The coding sequence may or may not contain a signal peptide or leader sequence. With the present invention, both the naturally occurring signal peptides or heterologous sequences can be used. Leader sequences can be removed by the host in post-translational processing. See, e.g., U.S. Patent Nos. 4,431,739; 4,425,437; 4,338,397. Such sequences include, but are not limited to, the TPA leader, as well as the honey bee mellitin signal sequence.

Other regulatory sequences may also be desirable which allow for regulation of expression of the protein sequences relative to the growth of the host cell. Such regulatory sequences are known to those of skill in the art, and examples include those which cause the expression of a gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Other types of regulatory elements may also be present in the vector, for example, enhancer sequences.

The control sequences and other regulatory sequences may be ligated to the coding sequence prior to insertion into a vector. Alternatively, the coding sequence can be cloned directly into an expression vector which already contains the control sequences and an appropriate restriction site.

In some cases it may be necessary to modify the coding sequence so that it may be attached to the control sequences with the appropriate orientation; i.e., to maintain the proper reading frame. Mutants or analogs may be prepared by the deletion of a portion of the sequence encoding the protein, by insertion of a sequence, and/or by substitution of one or more nucleotides within the sequence. Techniques for modifying nucleotide sequences, such as site-directed mutagenesis, are well known to those skilled in the art. See, e.g., Sambrook et al., *supra*; *DNA Cloning*, Vols. I and II, *supra*; *Nucleic Acid Hybridization*, *supra*.

The expression vector is then used to transform an appropriate host cell. A number of mammalian cell lines are known in the art and include immortalized cell lines available from the American Type Culture Collection (ATCC), such as, but not limited to, Chinese hamster ovary (CHO) cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), human hepatocellular carcinoma cells (e.g., Hep G2), Vero293 cells, as well as others. Similarly, bacterial hosts such as *E. coli*, *Bacillus subtilis*, and *Streptococcus spp.*, will find



use with the present expression constructs. Yeast hosts useful in the present invention include *inter alia*, *Saccharomyces cerevisiae*, *Candida albicans*, *Candida maltosa*, *Hansenula polymorpha*, *Kluyveromyces fragilis*, *Kluyveromyces lactis*, *Pichia guilliermondii*, *Pichia pastoris*, *Schizosaccharomyces pombe* and *Yarrowia lipolytica*. Insect cells for use  
5 with baculovirus expression vectors include, *inter alia*, *Aedes aegypti*, *Autographa californica*, *Bombyx mori*, *Drosophila melanogaster*, *Spodoptera frugiperda*, and *Trichoplusia ni*.

Depending on the expression system and host selected, the proteins of the present invention are produced by growing host cells transformed by an expression vector described  
10 above under conditions whereby the protein of interest is expressed. The selection of the appropriate growth conditions is within the skill of the art.

In one embodiment, the transformed cells secrete the polypeptide product into the surrounding media. Certain regulatory sequences can be included in the vector to enhance secretion of the protein product, for example using a tissue plasminogen activator (TPA)  
15 leader sequence, a  $\gamma$ -interferon signal sequence or other signal peptide sequences from known secretory proteins. The secreted polypeptide product can then be isolated by various techniques described herein, for example, using standard purification techniques such as but not limited to, hydroxyapatite resins, column chromatography, ion-exchange  
20 chromatography, size-exclusion chromatography, electrophoresis, HPLC, immunoabsorbent techniques, affinity chromatography, immunoprecipitation, and the like..

Alternatively, the transformed cells are disrupted, using chemical, physical or mechanical means, which lyse the cells yet keep the Env polypeptides substantially intact. Intracellular proteins can also be obtained by removing components from the cell wall or membrane, e.g., by the use of detergents or organic solvents, such that leakage of the Env  
25 polypeptides occurs. Such methods are known to those of skill in the art and are described in, e.g., *Protein Purification Applications: A Practical Approach*, (E.L.V. Harris and S. Angal, Eds., 1990)

For example, methods of disrupting cells for use with the present invention include but are not limited to: sonication or ultrasonication; agitation; liquid or solid extrusion; heat  
30 treatment; freeze-thaw; desiccation; explosive decompression; osmotic shock; treatment with lytic enzymes including proteases such as trypsin, neuraminidase and lysozyme; alkali treatment; and the use of detergents and solvents such as bile salts, sodium dodecylsulphate,

Triton, NP40 and CHAPS. The particular technique used to disrupt the cells is largely a matter of choice and will depend on the cell type in which the polypeptide is expressed, culture conditions and any pre-treatment used.

Following disruption of the cells, cellular debris is removed, generally by centrifugation, and the intracellularly produced Env polypeptides are further purified, using standard purification techniques such as but not limited to, column chromatography, ion-exchange chromatography, size-exclusion chromatography, electrophoresis, HPLC, immunoadsorbent techniques, affinity chromatography, immunoprecipitation, and the like.

For example, one method for obtaining the intracellular Env polypeptides of the present invention involves affinity purification, such as by immunoaffinity chromatography using anti-Env specific antibodies, or by lectin affinity chromatography. Particularly preferred lectin resins are those that recognize mannose moieties such as but not limited to resins derived from *Galanthus nivalis* agglutinin (GNA), *Lens culinaris* agglutinin (LCA or lentil lectin), *Pisum sativum* agglutinin (PSA or pea lectin), *Narcissus pseudonarcissus* agglutinin (NPA) and *Allium ursinum* agglutinin (AUA). The choice of a suitable affinity resin is within the skill in the art. After affinity purification, the Env polypeptides can be further purified using conventional techniques well known in the art, such as by any of the techniques described above.

It may be desirable to produce Env (e.g., gp120) complexes, either with itself or other proteins. Such complexes are readily produced by e.g., co-transfecting host cells with constructs encoding for the Env (e.g., gp120) and/or other polypeptides of the desired complex. Co-transfection can be accomplished either in *trans* or *cis*, i.e., by using separate vectors or by using a single vector which bears both of the Env and other gene. If done using a single vector, both genes can be driven by a single set of control elements or, alternatively, the genes can be present on the vector in individual expression cassettes, driven by individual control elements. Following expression, the proteins will spontaneously associate. Alternatively, the complexes can be formed by mixing the individual proteins together which have been produced separately, either in purified or semi-purified form, or even by mixing culture media in which host cells expressing the proteins, have been cultured. See, International Publication No. WO 96/04301, published February 15, 1996, for a description of such complexes.

Relatively small polypeptides, i.e., up to about 50 amino acids in length, can be conveniently synthesized chemically, for example by any of several techniques that are known to those skilled in the peptide art. In general, these methods employ the sequential addition of one or more amino acids to a growing peptide chain. Normally, either the amino or carboxyl group of the first amino acid is protected by a suitable protecting group. The protected or derivatized amino acid can then be either attached to an inert solid support or utilized in solution by adding the next amino acid in the sequence having the complementary (amino or carboxyl) group suitably protected, under conditions that allow for the formation of an amide linkage. The protecting group is then removed from the newly added amino acid residue and the next amino acid (suitably protected) is then added, and so forth. After the desired amino acids have been linked in the proper sequence, any remaining protecting groups (and any solid support, if solid phase synthesis techniques are used) are removed sequentially or concurrently, to render the final polypeptide. By simple modification of this general procedure, it is possible to add more than one amino acid at a time to a growing chain, for example, by coupling (under conditions which do not racemize chiral centers) a protected tripeptide with a properly protected dipeptide to form, after deprotection, a pentapeptide. See, e.g., J. M. Stewart and J. D. Young, Solid Phase Peptide Synthesis (Pierce Chemical Co., Rockford, IL 1984) and G. Barany and R. B. Merrifield, The Peptides: Analysis, Synthesis, Biology, editors E. Gross and J. Meienhofer, Vol. 2, (Academic Press, New York, 1980), pp. 3-254, for solid phase peptide synthesis techniques; and M. Bodansky, Principles of Peptide Synthesis, (Springer-Verlag, Berlin 1984) and E. Gross and J. Meienhofer, Eds., The Peptides: Analysis, Synthesis, Biology, Vol. 1, for classical solution synthesis.

Typical protecting groups include t-butyloxycarbonyl (Boc), 9-fluorenylmethoxycarbonyl (Fmoc) benzyloxycarbonyl (Cbz); p-toluenesulfonyl (Tx); 2,4-dinitrophenyl; benzyl (Bzl); biphenylisopropylloxycarboxy-carbonyl, t-amyloxycarbonyl, isobornyloxycarbonyl, o-bromobenzyloxycarbonyl, cyclohexyl, isopropyl, acetyl, o-nitrophenylsulfonyl and the like.

Typical solid supports are cross-linked polymeric supports. These can include divinylbenzene cross-linked-styrene-based polymers, for example, divinylbenzene-hydroxymethylstyrene copolymers, divinylbenzene-chloromethylstyrene copolymers and divinylbenzene-benzhydrylaminopolystyrene copolymers.

The polypeptide analogs of the present invention can also be chemically prepared by other methods such as by the method of simultaneous multiple peptide synthesis. See, e.g., Houghten *Proc. Natl. Acad. Sci. USA* (1985) 82:5131-5135; U.S. Patent No. 4,631,211.

## 5                    **Diagnostic and Vaccine Applications**

The intracellularly produced Env polypeptides of the present invention, complexes thereof, or the polynucleotides coding therefor, can be used for a number of diagnostic and therapeutic purposes. For example, the proteins and polynucleotides or antibodies generated against the same, can be used in a variety of assays, to determine the presence of reactive  
10    antibodies/and or Env proteins in a biological sample to aid in the diagnosis of HIV infection or disease status or as measure of response to immunization.

The presence of antibodies reactive with the Env (e.g., gp120) polypeptides and, conversely, antigens reactive with antibodies generated thereto, can be detected using standard electrophoretic and immunodiagnostic techniques, including immunoassays such as  
15    competition, direct reaction, or sandwich type assays. Such assays include, but are not limited to, western blots; agglutination tests; enzyme-labeled and mediated immunoassays, such as ELISAs; biotin/avidin type assays; radioimmunoassays; immunoelectrophoresis; immunoprecipitation, etc. The reactions generally include revealing labels such as fluorescent, chemiluminescent, radioactive, or enzymatic labels or dye molecules, or other  
20    methods for detecting the formation of a complex between the antigen and the antibody or antibodies reacted therewith.

Solid supports can be used in the assays such as nitrocellulose, in membrane or microtiter well form; polyvinylchloride, in sheets or microtiter wells; polystyrene latex, in beads or microtiter plates; polyvinylidene fluoride; diazotized paper; nylon membranes;  
25    activated beads, and the like.

Typically, the solid support is first reacted with the biological sample (or the gp120 proteins), washed and then the antibodies, (or a sample suspected of containing antibodies), applied. After washing to remove any non-bound ligand, a secondary binder moiety is added under suitable binding conditions, such that the secondary binder is capable of associating  
30    selectively with the bound ligand. The presence of the secondary binder can then be detected using techniques well known in the art. Typically, the secondary binder will comprise an antibody directed against the antibody ligands. A number of anti-human immunoglobulin

(Ig) molecules are known in the art (e.g., commercially available goat anti-human Ig or rabbit anti-human Ig). Ig molecules for use herein will preferably be of the IgG or IgA type, however, IgM may also be appropriate in some instances. The Ig molecules can be readily conjugated to a detectable enzyme label, such as horseradish peroxidase, glucose oxidase, Beta-galactosidase, alkaline phosphatase and urease, among others, using methods known to those of skill in the art. An appropriate enzyme substrate is then used to generate a detectable signal.

Alternatively, a "two antibody sandwich" assay can be used to detect the proteins of the present invention. In this technique, the solid support is reacted first with one or more of the antibodies directed against Env (e.g., gp120), washed and then exposed to the test sample. Antibodies are again added and the reaction visualized using either a direct color reaction or using a labeled second antibody, such as an anti-immunoglobulin labeled with horseradish peroxidase, alkaline phosphatase or urease.

Assays can also be conducted in solution, such that the viral proteins and antibodies thereto form complexes under precipitating conditions. The precipitated complexes can then be separated from the test sample, for example, by centrifugation. The reaction mixture can be analyzed to determine the presence or absence of antibody-antigen complexes using any of a number of standard methods, such as those immunodiagnostic methods described above.

The modified Env proteins, produced as described above, or antibodies to the proteins, can be provided in kits, with suitable instructions and other necessary reagents, in order to conduct immunoassays as described above. The kit can also contain, depending on the particular immunoassay used, suitable labels and other packaged reagents and materials (i.e. wash buffers and the like). Standard immunoassays, such as those described above, can be conducted using these kits.

The Env polypeptides and polynucleotides encoding the polypeptides can also be used in vaccine compositions, individually or in combination, in e.g., prophylactic (i.e., to prevent infection) or therapeutic (to treat HIV following infection) vaccines. The vaccines can comprise mixtures of one or more of the modified Env proteins (or nucleotide sequences encoding the proteins), such as Env (e.g., gp120) proteins derived from more than one viral isolate. The vaccine may also be administered in conjunction with other antigens and immunoregulatory agents, for example, immunoglobulins, cytokines, lymphokines, and chemokines, including but not limited to IL-2, modified IL-2 (cys125-ser125), GM-CSF, IL-

12,  $\gamma$ -interferon, IP-10, MIP1 $\beta$  and RANTES. The vaccines may be administered as polypeptides or, alternatively, as naked nucleic acid vaccines (*e.g.*, DNA), using viral vectors (*e.g.*, retroviral vectors, adenoviral vectors, adeno-associated viral vectors) or non-viral vectors (*e.g.*, liposomes, particles coated with nucleic acid or protein). The vaccines may also  
5 comprise a mixture of protein and nucleic acid, which in turn may be delivered using the same or different vehicles. The vaccine may be given more than once (*e.g.*, a "prime" administration followed by one or more "boosts") to achieve the desired effects. The same composition can be administered as the prime and as the one or more boosts. Alternatively, different compositions can be used for priming and boosting.

10 The vaccines will generally include one or more "pharmaceutically acceptable excipients or vehicles" such as water, saline, glycerol, ethanol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles.

A carrier is optionally present which is a molecule that does not itself induce the  
15 production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, lipid aggregates (such as oil droplets or liposomes), and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Furthermore, the Env  
20 polypeptide may be conjugated to a bacterial toxoid, such as toxoid from diphtheria, tetanus, cholera, etc.

Adjuvants may also be used to enhance the effectiveness of the vaccines. Such adjuvants include, but are not limited to: (1) aluminum salts (alum), such as aluminum hydroxide, aluminum phosphate, aluminum sulfate, etc.; (2) oil-in-water emulsion  
25 formulations (with or without other specific immunostimulating agents such as muramyl peptides (see below) or bacterial cell wall components), such as for example (a) MF59 (International Publication No. WO 90/14837), containing 5% Squalene, 0.5% Tween 80, and 0.5% Span 85 (optionally containing various amounts of MTP-PE (see below), although not required) formulated into submicron particles using a microfluidizer such as Model 110Y  
30 microfluidizer (Microfluidics, Newton, MA), (b) SAF, containing 10% Squalene, 0.4% Tween 80, 5% pluronic-blocked polymer L121, and thr-MDP (see below) either microfluidized into a submicron emulsion or vortexed to generate a larger particle size

emulsion, and (c) Ribi<sup>TM</sup> adjuvant system (RAS), (Ribi Immunochem, Hamilton, MT) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (Detox<sup>TM</sup>); (3) saponin adjuvants, such as Stimulon<sup>TM</sup> (Cambridge Bioscience, Worcester, MA) may be used or particle generated therefrom such as ISCOMs (immunostimulating complexes); (4) Complete Freund's Adjuvant (CFA) and Incomplete Freund's Adjuvant (IFA); (5) cytokines, such as interleukins (IL-1, IL-2, etc.), macrophage colony stimulating factor (M-CSF), tumor necrosis factor (TNF), etc.; (6) detoxified mutants of a bacterial ADP-ribosylating toxin such as a cholera toxin (CT), a pertussis toxin (PT), or an *E. coli* heat-labile toxin (LT), particularly LT-K63 (where lysine is substituted for the wild-type amino acid at position 63) LT-R72 (where arginine is substituted for the wild-type amino acid at position 72), CT-S109 (where serine is substituted for the wild-type amino acid at position 109), and PT-K9/G129 (where lysine is substituted for the wild-type amino acid at position 9 and glycine substituted at position 129) (see, e.g., International Publication Nos. W093/13202 and W092/19265); and (7) other substances that act as immunostimulating agents to enhance the effectiveness of the composition.

Muramyl peptides include, but are not limited to, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-normuramyl-L-alanyl-D-isoglutamine (nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine (MTP-PE), etc.

Typically, the vaccine compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. The preparation also may be emulsified or encapsulated in liposomes for enhanced adjuvant effect, as discussed above.

The vaccines will comprise a therapeutically effective amount of the modified Env proteins, or complexes of the proteins, or nucleotide sequences encoding the same, and any other of the above-mentioned components, as needed. By "therapeutically effective amount" is meant an amount of a modified Env (e.g., gp120) protein which will induce a protective immunological response in the uninfected, infected or unexposed individual to which it is administered. Such a response will generally result in the development in the subject of a secretory, cellular and/or antibody-mediated immune response to the vaccine. Usually, such

a response includes but is not limited to one or more of the following effects; the production of antibodies from any of the immunological classes, such as immunoglobulins A, D, E, G or M; the proliferation of B and T lymphocytes; the provision of activation, growth and differentiation signals to immunological cells; expansion of helper T cell, suppressor T cell, and/or cytotoxic T cell.

Preferably, the effective amount is sufficient to bring about treatment or prevention of disease symptoms. The exact amount necessary will vary depending on the subject being treated; the age and general condition of the individual to be treated; the capacity of the individual's immune system to synthesize antibodies; the degree of protection desired; the severity of the condition being treated; the particular Env polypeptide selected and its mode of administration, among other factors. An appropriate effective amount can be readily determined by one of skill in the art. A "therapeutically effective amount" will fall in a relatively broad range that can be determined through routine trials.

Once formulated, the nucleic acid vaccines may be accomplished with or without viral vectors, as described above, by injection using either a conventional syringe or a gene gun, such as the Accell® gene delivery system (PowderJect Technologies, Inc., Oxford, England). Delivery of DNA into cells of the epidermis is particularly preferred as this mode of administration provides access to skin-associated lymphoid cells and provides for a transient presence of DNA in the recipient. Both nucleic acids and/or peptides can be injected either subcutaneously, epidermally, intradermally, intramucosally such as nasally, rectally and vaginally, intraperitoneally, intravenously, orally or intramuscularly. Other modes of administration include oral and pulmonary administration, suppositories, needle-less injection, transcutaneous and transdermal applications. Dosage treatment may be a single dose schedule or a multiple dose schedule. Administration of nucleic acids may also be combined with administration of peptides or other substances.

While the invention has been described in conjunction with the preferred specific embodiments thereof, it is to be understood that the foregoing description as well as the examples which follow are intended to illustrate and not limit the scope of the invention. Other aspects, advantages and modifications within the scope of the invention will be apparent to those skilled in the art to which the invention pertains.



Experimental

Below are examples of specific embodiments for carrying out the present invention. The examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

- 5        Efforts have been made to ensure accuracy with respect to numbers used (e.g., amounts, temperatures, etc.), but some experimental error and deviation should, of course, be allowed for.

## EXAMPLE 1

10        A.1. Best-Fit and Homology Searches

The crystal structure of HXB-2 gp 120 was downloaded from the Brookhaven database (COMPLEX (HIV ENVELOPE PROTEIN/CD4/FAB) 15-JUN-98 1GC1 TITLE: HIV-1 GP120 CORE COMPLEXED WITH CD4 AND A NEUTRALIZING HUMAN ANTIBODY). Beta strands 3, 2, 21, and 20 of gp 120 form a sheet near the CD4 binding site. Strands  $\beta$ -3 and  $\beta$ -2 are connected by the V1/V2 loop. Strands  $\beta$ -21 and  $\beta$ -20 are connected by another small loop. The H-bonds at the interface between strands  $\beta$ -2 and  $\beta$ -21 are the only connection between domains of the "lower" half of the protein (joining helix alpha 1 to the CD4 binding site). This beta sheet and these loops mask some antigens (e.g., antigens which may generate neutralizing antibodies) that are only exposed during the CD4 binding.

20        Constructs that remove enough of the beta sheet to expose the antigens in the CD4 binding site, but leave enough of the protein to allow correct folding were designed. Specifically targeted were modifications to the small loop and, optional deletion of the V1/V2 loops. Three different types of constructs were designed: (1) constructs encoding polypeptides that leave the number of residues making up the entire 4-strand beta sheet intact, but replace one or more residues; (2) constructs that encode polypeptide having at least one residue of at least one beta strand excised or (3) constructs encoding polypeptides having at least two residues of at least one beta strand excised. Thus, a total of 6 different turns were needed to rejoin the ends of the strands.

30        Initially, residues in the small loop (residues 427-430, relative to HXB-2) and connected beta strands ( $\beta$ -20 and  $\beta$ -21) were modified to contain Gly and Pro (common in beta turns). These sequences were then used as the target to match in each search. The

geometry of the target was matched to known proteins in the Brookhaven Protein Data Bank. In particular, 5-residue turns (including an overlapping single residue at the N-terminal, the 2 residue target turn and 2 overlapping residues at the C-terminal) were searched in the databases. In other words, these modified loops add a 2 residue turn that should be able to support a geometry that will maintain the beta-sheet structure of the wild type protein. The calculations were performed using the default parameters in the Loop Search feature of the Biopolymer module of the Sybyl molecular modeling package. In each case, the 25 best fits based on geometry alone were reviewed and, of those, several selected for homology and fit.

In addition, it was also determined what modifications could be made to remove most of the V1/V2 loop (residues 124-198, relative to HXB-2) yet leave the geometry of the protein intact. As with the small loop, constructs were also designed which excised one or more residues from the  $\beta$ -2 strand (residues 119-123 of HXB-2), the  $\beta$ -3 strand (residues 199-201 of HXB-2) or both  $\beta$ -2 and  $\beta$ -3. For these constructs, known loops were searched to match the geometry of a pentamer (including two remaining residues from the N-terminal side, a 2 residue turn and 1 C-terminal residue). For these searches, Gly-Gly was preferred as the insert along with at least one C-terminal substitution.

#### A.2. Small Loop Replacements

In one aspect, the native sequence was replaced with residues that expose the CD4 binding site, but leave the overall geometry of the protein relatively unchanged. For the small loop replacements, the target to match was: ASN425-MET426-GLY427-GLY428-GLY431. Results of the search are summarized in Table 1.

Table 1: Search of Small Loop (Asn425 through Gly431)

Rank	Sequence	RMSD	% Homology	Seq Id No.
Best fit	LYS-ASP-SER-ASN-ASN	0.16689	62.5	27
3	TYR-GLY-LEU-GLY-LEU	0.220308	62.5	28
4	GLU-ARG-GLU-ASP-GLY	0.241754	62.5	29
7	ARG-LYS-GLY-GLY-ASN	0.24881	100	30
12	TRP-THR-GLY-SER-TYR	0.26417	83.33	31

Based on these results, constructs encoding Gly-Gly (#7), Gly-Ser (#12) or Gly-Gly-Asn (#7) were recommended.

- As V1/V2 and one or more residues of  $\beta$ -2 and  $\beta$ -3 are also optionally deleted in the modified polypeptides of the invention, known loops to match the geometry of the V1/V2 loop were also searched. The V1/V2 loop the target to match was: Lys121-Leu-122-Gly123-Gly124-Ser199. Some notable matches are shown in Table 2:

Table 2: Search of V1/V2 loop (Lys121 through Ser199)

Rank	Sequence	RMSD	% Homology	Seq Id. No.
Best fit	GLN-VAL-HIS-ASP-GLU	0.154764	68.75	32
2	LYS-GLU-GLY-ASP-LYS	0.15718	81.25	33
9	ARG-SER-GLY-ARG-SER	0.173731	68.75	34
11	THR-LEU-GLY-ASN-SER	0.175554	81.25	35
16	HIS-PHE-GLY-ALA-GLY	0.178772	93.75	36

Based on these searches, constructs encoding Gly-Asn in place of V1/V2 were recommended.

### A.3. One Additional Residue Excisions

For a slightly truncated small loop, one more residue was trimmed from each beta strand to slightly shorten the beta sheet. The target to match was: ILE424-ASN425-GLY426-GLY427-LYS432. Results are shown in Table 3:

Table 3: Search of Beta sheet shortened by One residue (Ile424 through Lys432)

Rank	Sequence	RMSD	% Homology	Seq Id No.
Best fit:	ARG-MET-ALA-PRO-VAL	0.316805	58.33	37
Best hom:	ASP-SER-ASP-GLY-PRO	0.440896	83.33	38

Although these searches showed more variation and worse fits than the previous truncation, the Pro-Val or Pro-Leu encoding constructs were very similar. Accordingly, Ala-Pro encoding constructs were recommended.

Sequences encoding gp120 polypeptides having V1/V2 deleted and an additional  
 5 residue from  $\beta$ -2 or  $\beta$ -3 excised were also searched. The V1/V2 loop the target to match was:  
 VAL120-LYS121-GLY122-GLY123-VAL200. Some notable matches are shown in Table 4.

Table 4: Search of V1/V2 loop (Val120 through Val200)

10	Rank	Sequence	RMSD	% Homology	Seq Id No
	Best fit:	THR-VAL-ASP-PRO-TYR	0.400892	58.33333	39
	2	SER-THR-ASN-PRO-LEU	0.402575	54.16667	40
	3	THR-ARG-SER-PRO-LEU	0.403965	58.33333	41
15	7	ARG-MET-ALA-PRO-VAL	0.440118	58.33333	42

The construct encoding Ala-Pro (e.g., #7) was recommended.

#### A.4. Further Excisions

In yet another truncation, an additional residue was trimmed from the  $\beta$ -20 and  $\beta$ -21  
 20 strands to further shorten the beta sheet. The target to match was ILE423-ILE424-GLY425-  
 GLY426-ALA433. Notable matches are shown in Table 5.

Table 5: Search of Beta sheet shortened by Two Residues (Ile423 through Ala433)

25	Rank	Sequence	RMSD	% Homology	Seq Id No
	Best fit:	THR-TYR-GLU-GLY-VAL	0.130107	79.16666	43
	2	GLN-VAL-GLY-ASN-THR	0.138245	79.16666	44
	3:	THR-VAL-GLY-GLY-ILE	0.153362	100	45

A construct encoding Gly-Gly (e.g., #3), which has 100% homology, was  
 30 recommended.

Also searched were sequences encoding a deleted V1/V2 region and at least two residues excised from  $\beta$ -2,  $\beta$ -3 or at least one residue excised from  $\beta$ -2 and  $\beta$ -3. The target to match was: CYS119-VAL120-GLY121-GLY122-ILE201. Notable matches are shown in Table 6.

Table 6: Search of V1/V2 loop (Cys119 through Ile201)

Rank	Sequence	RMSD	% Homology	Seq Id No
Best fit:	ASP-LEU-PRO-GLY-CYS	0.250501	75	46
4	ASP-VAL-GLY-GLY-LEU	0.290383	100	47

It was determined that both constructs would be used.

#### B.1. Constructs encoding modified Env polypeptides

As described above, the native loops extruding from the 4- $\beta$  antiparallel-stands were excised and replaced with 1 to 3 residue turns. The loops were replaced so as to leave the entire  $\beta$ -strands or excised by trimming one or more amino acid from each side of the connected strands. The ends of the strands were rejoined with turns that preserve the same backbone geometry (e.g., tertiary structure of  $\beta$ -20 and  $\beta$ -21), as determined by searching the Brookhaven Protein Data Bank.

Table 7A is a summary of the truncations of the variable regions 1 and 2 recommended for this study, as determined in Example 1.A. above.

Table 7A

V1/V2 Modifications	SEQ ID NO	Figure
-LEU122-GLY-ASN-SER199	7	10
-LYS121-ALA-PRO-VAL200-	6	9
-VAL120-GLY-GLY-ILE201-	4	7
-VAL120-PRO-GLY-ILE201B-	5	8
-VAL120-GLY-ALA-GLY-ALA204-	3	6
-VAL120-GLY-GLY-ALA-THR202-	8	11
-VAL127-GLY-ALA-GLY-ASN195-	25	28

As previously noted, the polypeptides encoded by the constructs of the present invention are numbered relative to HXB-2, but the particular amino acid residue of the polypeptides encoded by these exemplary constructs is based on SF-162. Thus, for example, although amino acid residue 195 in HXB-2 is a serine (S), constructs encoding polypeptides having then wild type SF162 sequence will have an asparagine (N) at this position. Table 7B shows just three of the variations in amino acid sequence between strains HXB-2 and SF162. The entire sequences, including differences in residue and amino acid number, of HXB-2 and SF162 are shown in the alignment of Figure 2 (SEQ ID NOs:1 and 2).

Table 7B

HXB-2 amino acid number	HXB-2 Residue	SF162 Residue/amino acid number
128	Serine (S)	Thr (T)/114
195	Serine (S)	Asn (N)/188
426	Met (M)	Arg (R)/411

Constructs containing deletions in the  $\beta$ -20 strand,  $\beta$ -21 stand and small loop were also constructed. Shown in Table 8 are constructs encoding truncations in these regions. The constructs in Table 8 are numbered relative to HXB-2 but the unmodified amino acid sequence is based on SF162. Thus, the construct encodes an arginine (Arg) as is found in

SF162 in the amino acid position numbered 426 relative to HXB-2 (See, also, Table 7B). Changes from wildtype (SF162) are shown in bold in Table 8B.

Table 8

Small Loop/ $\beta$ -20 and $\beta$ -21 (Modified)	SEQ ID NO	Figure
-TRP427- <b>GLY</b> -GLY431-	9	12
-ARG426- <b>GLY-GLY</b> -GLY431-	10	13
-ARG426- <b>GLY-SER</b> -GLY431B-	11	14
-ARG426- <b>GLY-GLY</b> -ASN-LYS432-	12	15
-ASN425- <b>ALA-PRO</b> -LYS432-	13	16
-ILE424- <b>GLY-GLY</b> -ALA433-	14	17
-ILE423- <b>GLY-GLY</b> -MET434-	15	18
GLN422- <b>GLY-GLY</b> -TYR435-	16	19
-GLN422- <b>ALA-PRO</b> -TYR435B-	17	20

The deletion constructs shown in Tables 7 and 8 for each one of the  $\beta$ -strands and combinations of them are constructed. These deletions will be tested in the Env forms gp120, gp140 and gp160 from different HIV strains like subtype B strains (e.g., SF162, US4, SF2), subtype E strains (e.g., CM235) and subtype C strains (e.g., AF110968 or AF110975).

Exemplary constructs for SF162 are shown in the

Figures and are summarized in Table 9. As noted above in Figure 2 and Table 7B, in the bridging sheet region, the amino acid sequence of SF162 differs from HXB-2 in that the Met426 of HXB-2 is an Arg in SF162. In Table 9, V1/V2 refers to deletions in the V1/V2 region; # bsm refers to a modification in the bridging sheet small loop.

Table 9

Construct	Seq. Id.	Fig.	Modification/Amino acid sequence
Val120-Ala204	3	6	V1/V2: Val120-Gly-Ala-Gly-Ala204
Val120-Ile201	4	7	V1/V2: Val120-Gly-Gly-Ile201
Val120-Ile201B	5	8	V1/V2: Val120- <b>Pro</b> -Gly-Ile201
Lys121-Val200	6	9	V1/V2: Lys121- <b>Ala-Pro</b> -Val200

Table 9

Construct	Seq. Id.	Fig.	Modification/Amino acid sequence
Leu122-Ser199	7	10	V1/V2: Leu122-Gly-Asn-Ser199
Val120-Thr202	8	11	V1/V2: Val120-Gly-Gly-Ala-Thr202
Trp427-Gly431	9	12	bsm: Trp427-Gly-Gly431
Arg426-Gly431	10	13	bsm: Arg426-Gly-Gly-Gly431
Arg426-Gly431B	11	14	bsm: Arg426-Gly-Ser-Gly431
Arg426-Lys432	12	15	bsm: Arg426-Gly-Gly-Asn-Lys432
Asn425-Lys432	13	16	bsm: Asn425-Ala-Pro-Lys432
Ile424-Ala433	14	17	bsm: Ile424-Gly-Gly-Ala433
Ile423-Met434	15	18	bsm: Ile423-Gly-Gly-Met434
Gln422-Tyr435	16	19	bsm: Gln422-Gly-Gly-Tyr435
Val127-Asn195	25	28	bsm: Val127-Gly-Ala-Gly-Asn195
Gln422-Tyr435B	17	20	bsm: Gln422-Ala-Pro-Tyr435
Leu122-Ser199; Arg426-Gly431	18	21	V1/V2/bsm: Leu122-Gly-Asn-Ser199 --- Arg426-Gly-Gly-Gly431
Leu122-Ser199; Arg426-Lys432	19	22	V1/V2/bsm: Leu122-Gly-Asn-Ser199 --- Arg426-Gly-Gly-Asn-Lys432
Leu122-Ser199-Trp427-Gly431	20	23	V1/V2/bsm: Leu122-Gly-Asn-Ser199 --- Trp427-Gly-Gly431
Lys121-Val200-Asn425-Lys432	21	24	V1/V2/bsm: Lys121-Ala-Pro-Val200 --- Asn425-Ala-Pro-Lys432
Val120-Ile201-Ile424-Ala433	22	25	V1/V2/bsm: Val120-Gly-Gly-Ile201 --- Ile424-Gly-Gly-Ala433
Val120-Ile201B-Ile424-Ala433	23	26	V1/V2/bsm: Val120-Pro-Gly-Ile201 --- Ile424-Gly-Gly-Ala433
Val120-Thr202; Ile424-Ala433	24	27	V1/V2/bsm: Val120-Gly-Gly-Ala-Thr202 --- Ile424-Gly-Gly-Ala433
Val127-Asn195; Arg426-Gly431	25	29	V1/V2/bsm: Val127-Gly-Ala-Gly-Asn195 --- Arg426-Gly-Gly-Gly431

Combinations of V1/V2 deletions and bridging sheet small loop modifications in addition to those specifically shown in Table 9 are also within the scope of the present invention. Various forms of the different embodiments of the invention, described herein, may be combined.



The first screening will be done after transient expression in COS-7, RD and/or 293 cells. The proteins that are expressed will be analyzed by immunoblot, ELISA, and for binding to mAbs directed to the CD4 binding site and other important epitopes on gp120 to determine integrity of structure. They will also be tested in a CD4 binding assay and, in  
5 addition, the binding of neutralizing antibodies, for example using patient sera or mAb 448D (directed to Glu370 and Tyr384, a region of the CD4 binding groove that is not altered by the deletions).

The immunogenicity of these novel Env glycoproteins will be tested in rodents and primates. The structures will be administered as DNA vaccines or adjuvanted protein  
10 vaccines or in combined modalities. The goal of these vaccinations will be to archive broadly reactive neutralizing antibody responses.

Claims:

What is claimed is:

- 5           1. A polynucleotide encoding a modified HIV Env polypeptide wherein the polypeptide has at least one amino acid deleted or replaced in the region corresponding to residues 420 to 436 relative to HXB-2 (SEQ ID NO:1).
2. The polynucleotide of claim 1, wherein the region corresponding to residues 124-  
10   198 relative to HXB-2 is deleted and at least one amino acid is deleted or replaced in the regions corresponding to the residues 119 to 123 and 199 to 210 relative to HXB-2 (SEQ ID NO:1).
3. The polynucleotide of claim 1, wherein at least one amino acid in the region  
15   corresponding to residues 427 through 429 relative to HXB-2 (SEQ ID NO:1) is deleted or replaced.
4. The polynucleotide of claim 2, wherein at least one amino acid of the in the region  
20   corresponding to residues 427 through 429 relative to HXB-2 (SEQ ID NO:1) is deleted or replaced.
5. The polynucleotide of claim 1, wherein the amino acid sequence of the modified HIV Env polypeptide is based on strain SF162.
- 25           6. An immunogenic modified HIV Env polypeptide having at least one amino acid deleted or replaced in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).
7. The polypeptide of claim 6, wherein one amino acid is deleted in the region  
30   corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).

8. The polypeptide of claim 6, wherein more than one amino acid is deleted in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).

5 9. The polypeptide of claim 6, wherein at least one amino acid is replaced in the region corresponding to residues 420 through 436, relative to HXB-2 (SEQ ID NO:1).

10 10. The polypeptide of claim 6, wherein at least one amino acid residue between amino acid residue 427 and amino acid residue 429 relative to HXB-2 (SEQ ID NO:1) is deleted or replaced.

11. The polypeptide of claim 6, wherein the V1 and V2 regions of the polypeptide are truncated.

15 12. The polypeptide of claim 10, wherein the V1 and V2 regions of the polypeptide are truncated.

13. The polypeptide of claim 6, wherein the amino acid sequence of the modified HIV Env polypeptide is based on strain SF162.

20 14. A construct comprising the nucleotide sequence depicted in Figure 6 (SEQ ID NO:3).

25 15. A construct comprising the nucleotide sequence depicted in Figure 7 (SEQ ID NO:4).

16. A construct comprising the nucleotide sequence depicted in Figure 8 (SEQ ID NO:5).

30 17. A construct comprising the nucleotide sequence depicted in Figure 9 (SEQ ID NO:6).

18. A construct comprising the nucleotide sequence depicted in Figure 10 (SEQ ID NO:7).

19. A construct comprising the nucleotide sequence depicted in Figure 11 (SEQ ID NO:8).

20. A construct comprising the nucleotide sequence depicted in Figure 12 (SEQ ID NO:9).

21. A construct comprising the nucleotide sequence depicted in Figure 13 (SEQ ID NO:10).

22. A construct comprising the nucleotide sequence depicted in Figure 14 (SEQ ID NO:11).

23. A construct comprising the nucleotide sequence depicted in Figure 15 (SEQ ID NO:12).

24. A construct comprising the nucleotide sequence depicted in Figure 16 (SEQ ID NO:13).

25. A construct comprising the nucleotide sequence depicted in Figure 17 (SEQ ID NO:14).

26. A construct comprising the nucleotide sequence depicted in Figure 18 (SEQ ID NO:15).

27. A construct comprising the nucleotide sequence depicted in Figure 19 (SEQ ID NO:16).

28. A construct comprising the nucleotide sequence depicted in Figure 20 (SEQ ID NO:17).

29. A construct comprising the nucleotide sequence depicted in Figure 21 (SEQ ID NO:18).

5 30. A construct comprising the nucleotide sequence depicted in Figure 22 (SEQ ID NO:19).

31. A construct comprising the nucleotide sequence depicted in Figure 23 (SEQ ID NO:20).

10 32. A construct comprising the nucleotide sequence depicted in Figure 24 (SEQ ID NO:21).

33. A construct comprising the nucleotide sequence depicted in Figure 25 (SEQ ID NO:22).

15 34. A construct comprising the nucleotide sequence depicted in Figure 26 (SEQ ID NO:23).

20 35. A construct comprising the nucleotide sequence depicted in Figure 27 (SEQ ID NO:24).

36. A construct comprising the nucleotide sequence depicted in Figure 28 (SEQ ID NO:25).

25 37. A construct comprising the nucleotide sequence depicted in Figure 29 (SEQ ID NO:26).

38. A vaccine composition comprising a polynucleotide encoding a modified Env polypeptide according to any one of claims 1-5.

30 39. A vaccine composition comprising a polynucleotide construct encoding a modified Env polypeptide according to any of claims 14-37.

40. A vaccine composition comprising a modified Env polypeptide according to any of claims 6-13.

5 41. The vaccine composition of any of claims 38-40, further comprising an adjuvant.

42. A method of inducing an immune response in subject comprising, administering a polynucleotide according to any one of claims 1-5 in an amount sufficient to induce an immune response in the subject.

10 43. A method of inducing an immune response in subject comprising, administering a polynucleotide construct according to any one of claims 14-37 in an amount sufficient to induce an immune response in the subject.

15 44. A method of inducing an immune response in a subject comprising administering a composition comprising a modified Env polypeptide according to any one of claims 6-13, wherein the composition is administered in an amount sufficient to induce an immune response in the subject

20 45. The method of any of claims 42-44 further comprising administering an adjuvant to the subject.

46. A method of inducing an immune response in a subject comprising  
(a) administering a first composition comprising a polynucleotide according to any of claims 1-5 in a priming step and  
25 (b) administering a second composition comprising a modified Env polypeptide according to any of claims 6-13, as a booster, in an amount sufficient to induce an immune response in the subject.

30 47. The method of claim 46 wherein the first composition or second composition further comprise an adjuvant.

48. The method of claim 46 wherein the first and second compositions further comprise an adjuvant.

# gp120 core structure

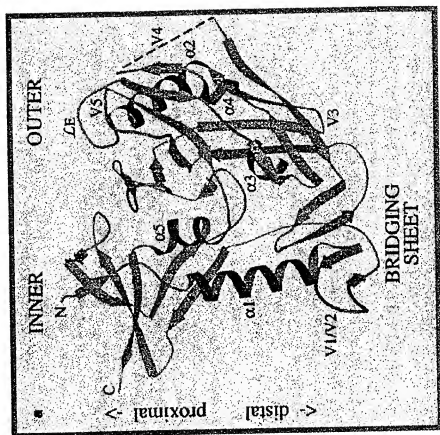


FIG. 1



FIG. 2A

		351	*			400
HXB2	(323)	TIGNDIRQAHCHNISRAKWNNTLKQIASLREQFGNNKTIIFNQSSGGDPEI				
162	(314)	TIGDIEKHCHNISGEKNNNTKQIVTQLQAOEG-NWATFEKQSSGGDPEI				
SF2	(324)	TIGDIEKHCHNISQAQNNNTEDIVKLRQEGNNKTIIFNQSSGGDPEI				
CM236	(324)	TIGDIEKHCHNISQNGTKNEVLTQVTEKHEH-NWATFEKQSSGGDPEI				
US4	(334)	TIGDIEKHCHNISQANNNTEDIVKLRQEGNNKTIIFNQSSGGDPEI				
Consensus	(351)	TIGDIRQAHCHNISRAKWNNTLQIVKLREQFGNNKTIIFNQSSGGDPEI				
		401	*	*		450
HXB2	(372)	VTHSNGGEGFSTSTQSTSTWTFNSTWSTFEGSNNTGEGSDITITPDRBK				
162	(363)	VMHSNGGEGFSTSTQSTSTWTFNSTWTFEGSNNTGEGSDITITPDRBK				
SF2	(374)	VMHSNGGEGFSTSTQSTSTWTFNSTWTFEGSNNTGEGSDITITPDRBK				
CM236	(373)	TMHSNGGEGFSTSTQSTSTWTFNSTWTFEGSNNTGEGSDITITPDRBK				
US4	(384)	VTHSNGGEGFSTSTQSTSTWTFNSTWTFEGSNNTGEGSDITITPDRBK				
Consensus	(401)	VMHSFNCGGEFFYCNTTQLENSTWNTGNTGTITILPCRIK				
		↓		↓		
		451		*		500
HXB2	(422)	QIINHWQEVGKAMYAPPIGQIRCSNITGLLLTRDGGNITNDTEFF				
162	(407)	QIINHWQEVGKAMYAPPIGQIRCSNITGLLLTRDGGNITNDTEFF				
SF2	(419)	QIINHWQEVGKAMYAPPIGQIRCSNITGLLLTRDGGNITNDTEFF				
CM236	(417)	QIINHWQEVGKAMYAPPIGQIRCSNITGLLLTRDGGNITNDTEFF				
US4	(430)	QIINHWQEVGKAMYAPPIGQIRCSNITGLLLTRDGGNITNDTEFF				
Consensus	(451)	QIINHWQEVGKAMYAPPIGQIRCSNITGLLLTRDGGNITNDTEFF				
		501		*		550
HXB2	(469)	REGGDMRDNRSELYKYKVVKIEPLGVAPTKAKRRVQREKRAVGI				
162	(455)	REGGDMRDNRSELYKYKVVKIEPLGVAPTKAKRRVQREKRAVGI				
SF2	(467)	REGGDMRDNRSELYKYKVVKIEPLGVAPTKAKRRVQREKRAVGI				
CM236	(464)	REGGDMRDNRSELYKYKVVKIEPLGVAPTKAKRRVQREKRAVGI				
US4	(480)	REGGDMRDNRSELYKYKVVKIEPLGVAPTKAKRRVQREKRAVGI				
Consensus	(501)	REGGDMRDNRSELYKYKVVKIEPLGVAPTKAKRRVQREKRAVGI				
		551		*		600
HXB2	(518)	MFLGFLGAAGSTMGAASLTLTVQARQLLSGTVQQNNLLRATEAQOQLLQ				
162	(504)	MFLGFLGAAGSTMGAASLTLTVQARQLLSGTVQQNNLLRATEAQOQLLQ				
SF2	(517)	MFLGFLGAAGSTMGAASLTLTVQARQLLSGTVQQNNLLRATEAQOQLLQ				
CM236	(513)	MFLGFLGAAGSTMGAASLTLTVQARQLLSGTVQQNNLLRATEAQOQLLQ				
US4	(529)	MFLGFLGAAGSTMGAASLTLTVQARQLLSGTVQQNNLLRATEAQOQLLQ				
Consensus	(551)	MFLGFLGAAGSTMGAASLTLTVQARQLLSGTVQQNNLLRATEAQOQLLQ				
		601		*	*	650
HXB2	(568)	LTVWGIKQLQARVLAVERYLKQOQLGIGWCSGKLICTTAVPWNASWSNK				
162	(554)	LTVWGIKQLQARVLAVERYLKQOQLGIGWCSGKLICTTAVPWNASWSNK				
SF2	(567)	LTVWGIKQLQARVLAVERYLKQOQLGIGWCSGKLICTTAVPWNASWSNK				
CM236	(563)	LTVWGIKQLQARVLAVERYLKQOQLGIGWCSGKLICTTAVPWNASWSNK				
US4	(579)	LTVWGIKQLQARVLAVERYLKQOQLGIGWCSGKLICTTAVPWNASWSNK				
Consensus	(601)	LTVWGIKQLQARVLAVERYLKQOQLGIGWCSGKLICTTAVPWNASWSNK				

FIG. 2B

	651		700
HXB2 (618)	SLDQNNHTTWMEERDRNNNTSLCHSLIEESQNOQEKNEQB	GGSHKQWA	
162	SLDQNNMTTWMEERDRNNNTSLCHSLIEESQNOQEKNEQB	GGSHKQWA	
SF2 (617)	SLDQNNMTTWMEERDRNNNTSLCHSLIEESQNOQEKNEQB	GGSHKQWA	
CM236 (613)	SYEYNNMTTWMEERDRNNNTSLCHSLIEESQNOQEKNEQB	GGSHKQWA	
US4 (629)	SLTNNMTTWMEERDRNNNTSLCHSLIEESQNOQEKNEQB	GGSHKQWA	
Consensus (651)	SLEETIWNMTTWMEWEREI	NYTNLIYTLIEESQNOQEKNEQELLELDKWA	
	701		750
HXB2 (668)	SANNRNTNTNNAKPKKAKLVGSLVSRVVA	VSQVNRVROHSHLSF	
162 (654)	SANNRDRSKKPKPKKI	SLSLVSRVVA	VSQVNRVROHSHLSF
SF2 (667)	SANNRNTNTNNAKPKKAKLVGSLVSRVVA	VSQVNRVROHSHLSF	
CM236 (663)	SANNRDRSKKPKPKKI	SLSLVSRVVA	VSQVNRVROHSHLSF
US4 (679)	SANNRNTNTNNAKPKKAKLVGSLVSRVVA	VSQVNRVROHSHLSF	
Consensus (701)	SLANNRNTNTNNAKPKKAKLVGSLVSRVVA	VSQVNRVROHSHLSF	
	751		*800
HXB2 (718)	QHHLPTPEGGRPEGEGGERDRDRSRVLNDS	LALITDDDRNRHGS	
162 (704)	QERFPAPGGRPEGEGGERDRDRSRVLSL	LALITDDDRNRHGS	
SF2 (717)	QRLPVPEGGRPDGEGGERDRDRSRVLR	QFSLITGSHLSLGS	
CM236 (713)	QPFPHQEGGRSEREGGSGQGRDRSRVLR	QFSLITGSHLSLGS	
US4 (729)	QRLPAQEGGRPEGEGGERDRDRSRVLR	LALITDDDRNRHGS	
Consensus (751)	QTRLPGRGPDPRGEGIEEGGERDRDRSRVLV	G LALIWDDLRSCLFS	
	801		850
HXB2 (768)	YHHRDLLAEVATVEIQLGR-----	RGWEALKYWNWNLQWSDHAKNS	
162 (754)	YHHRDLLAEVATVEIQLGR-----	RGWEALKYWNWNLQWSDHAKNS	
SF2 (767)	YHHRDLLAEVATVEIQLGR-----	RGWEALKYWNWNLQWSDHAKNS	
CM236 (763)	YHHRDLLAEVATVEIQLGR-----	RGWEALKYWNWNLQWSDHAKNS	
US4 (779)	YHHRDLLAEVATVEIQLGR-----	RGWEALKYWNWNLQWSDHAKNS	
Consensus (801)	YHRLRDLLEIAARIVELLGR	RGWEALKYWNWNLQW QELKNS	
	851		900
HXB2 (811)	AVSLLNATAVAEGTORVIEVQRAFRAILHIPRRIRQGLER	LL----	
162 (797)	AVSLFDATAVAEGTORVIEVQRAFRAILHIPRRIRQGLER	LL----	
SF2 (810)	AVSLNATAVAEGTORVIEVQRAFRAILHIPRRIRQGLER	LL----	
CM236 (813)	AVSLLDATAVAEGTORVIEVQRAFRAILHIPRRIRQGLER	LL----	
US4 (822)	AVSLNATAVAEGTORVIEVQRAFRAILHIPRRIRQGLER	LL----	
Consensus (851)	AVSLNATAVAEGTORVIEVQRAFRAILHIPRRIRQGLER	LL	

FIG. 2C

		1	40
Leu122-Ser199	(1)	GAATTCGCCACCATTGGATGCAATGAAGAGAGGGCTCTGCT	
Val1127-Asn195	(1)	GAATTCGCCACCATTGGATGCAATGAAGAGAGGGCTCTGCT	
Val1120-11e201B	(1)	GAATTCGCCACCATTGGATGCAATGAAGAGAGGGCTCTGCT	
Val1120-Ala204	(1)	GAATTCGCCACCATTGGATGCAATGAAGAGAGGGCTCTGCT	
Val1120-11e201	(1)	GAATTCGCCACCATTGGATGCAATGAAGAGAGGGCTCTGCT	
Val1120-Thr202	(1)	GAATTCGCCACCATTGGATGCAATGAAGAGAGGGCTCTGCT	
Lys121-Val200	(1)	GAATTCGCCACCATTGGATGCAATGAAGAGAGGGCTCTGCT	
Consensus	(1)	GAATTCGCCACCATTGGATGCAATGAAGAGAGGGCTCTGCT	
		41	80
Leu122-Ser199	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCCAG	
Val1127-Asn195	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCCAG	
Val1120-11e201B	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCCAG	
Val1120-Ala204	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCCAG	
Val1120-11e201	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCCAG	
Val1120-Thr202	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCCAG	
Lys121-Val200	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCCAG	
Consensus	(41)	GTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCCAG	
		81	120
Leu122-Ser199	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGACTACGGCGTG	
Val1127-Asn195	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGACTACGGCGTG	
Val1120-11e201B	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGACTACGGCGTG	
Val1120-Ala204	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGACTACGGCGTG	
Val1120-11e201	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGACTACGGCGTG	
Val1120-Thr202	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGACTACGGCGTG	
Lys121-Val200	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGACTACGGCGTG	
Consensus	(81)	CGCCGTGGAGAAGCTGTGGGTGACCGTGACTACGGCGTG	
		121	160
Leu122-Ser199	(121)	CCCGTGTGGAAGGAGGCCACCAACCCCTGTTCTGCGGCA	
Val1127-Asn195	(121)	CCCGTGTGGAAGGAGGCCACCAACCCCTGTTCTGCGGCA	
Val1120-11e201B	(121)	CCCGTGTGGAAGGAGGCCACCAACCCCTGTTCTGCGGCA	
Val1120-Ala204	(121)	CCCGTGTGGAAGGAGGCCACCAACCCCTGTTCTGCGGCA	
Val1120-11e201	(121)	CCCGTGTGGAAGGAGGCCACCAACCCCTGTTCTGCGGCA	
Val1120-Thr202	(121)	CCCGTGTGGAAGGAGGCCACCAACCCCTGTTCTGCGGCA	
Lys121-Val200	(121)	CCCGTGTGGAAGGAGGCCACCAACCCCTGTTCTGCGGCA	
Consensus	(121)	CCCGTGTGGAAGGAGGCCACCAACCCCTGTTCTGCGGCA	
		161	200
Leu122-Ser199	(161)	GCGACGCCAAGGCTTACGACACCGAGGTGCACAACGTGTG	
Val1127-Asn195	(161)	GCGACGCCAAGGCTTACGACACCGAGGTGCACAACGTGTG	
Val1120-11e201B	(161)	GCGACGCCAAGGCTTACGACACCGAGGTGCACAACGTGTG	
Val1120-Ala204	(161)	GCGACGCCAAGGCTTACGACACCGAGGTGCACAACGTGTG	
Val1120-11e201	(161)	GCGACGCCAAGGCTTACGACACCGAGGTGCACAACGTGTG	
Val1120-Thr202	(161)	GCGACGCCAAGGCTTACGACACCGAGGTGCACAACGTGTG	
Lys121-Val200	(161)	GCGACGCCAAGGCTTACGACACCGAGGTGCACAACGTGTG	
Consensus	(161)	GCGACGCCAAGGCTTACGACACCGAGGTGCACAACGTGTG	
		201	240
Leu122-Ser199	(201)	GGCCACCCACGCTTCGTTGCCACCGAACCCCAACCCCCAG	
Val1127-Asn195	(201)	GGCCACCCACGCTTCGTTGCCACCGAACCCCAACCCCCAG	
Val1120-11e201B	(201)	GGCCACCCACGCTTCGTTGCCACCGAACCCCAACCCCCAG	
Val1120-Ala204	(201)	GGCCACCCACGCTTCGTTGCCACCGAACCCCAACCCCCAG	
Val1120-11e201	(201)	GGCCACCCACGCTTCGTTGCCACCGAACCCCAACCCCCAG	
Val1120-Thr202	(201)	GGCCACCCACGCTTCGTTGCCACCGAACCCCAACCCCCAG	
Lys121-Val200	(201)	GGCCACCCACGCTTCGTTGCCACCGAACCCCAACCCCCAG	
Consensus	(201)	GGCCACCCACGCTTCGTTGCCACCGAACCCCAACCCCCAG	
		241	280
Leu122-Ser199	(241)	GAGATCGTGTGGAGAAGCTGACCGAGAAGTCTCAACATGT	
Val1127-Asn195	(241)	GAGATCGTGTGGAGAAGCTGACCGAGAAGTCTCAACATGT	

FIG. 3A

Vall120-11e201B	(241)	GAGATCGTGTGGAGAACGTGACCGAGAACCTTCAACATGT	
Vall120-Ala204	(241)	GAGATCGTGTGGAGAACGTGACCGAGAACCTTCAACATGT	
Vall120-11e201	(241)	GAGATCGTGTGGAGAACGTGACCGAGAACCTTCAACATGT	
Vall120-Thr202	(241)	GAGATCGTGTGGAGAACGTGACCGAGAACCTTCAACATGT	
Lys121-Val200	(241)	GAGATCGTGTGGAGAACGTGACCGAGAACCTTCAACATGT	
Consensus	(241)	GAGATCGTGTGGAGAACGTGACCGAGAACCTTCAACATGT	281 320
Leu122-Ser199	(281)	GGAGAACAACATGGTGGAGCAGATGCACGAGGACATCAT	
Vall127-Asn195	(281)	GGAGAACAACATGGTGGAGCAGATGCACGAGGACATCAT	
Vall120-11e201B	(281)	GGAGAACAACATGGTGGAGCAGATGCACGAGGACATCAT	
Vall120-Ala204	(281)	GGAGAACAACATGGTGGAGCAGATGCACGAGGACATCAT	
Vall120-11e201	(281)	GGAGAACAACATGGTGGAGCAGATGCACGAGGACATCAT	
Vall120-Thr202	(281)	GGAGAACAACATGGTGGAGCAGATGCACGAGGACATCAT	
Lys121-Val200	(281)	GGAGAACAACATGGTGGAGCAGATGCACGAGGACATCAT	
Consensus	(281)	GGAGAACAACATGGTGGAGCAGATGCACGAGGACATCAT	321 360
Leu122-Ser199	(321)	CAGCCTGTGGGACCAAGCCTGAAGCCCTGCGTGAAGCTG	
Vall127-Asn195	(321)	CAGCCTGTGGGACCAAGCCTGAAGCCCTGCGTGAAGCTG	
Vall120-11e201B	(321)	CAGCCTGTGGGACCAAGCCTGAAGCCCTGCGTGGC----	
Vall120-Ala204	(321)	CAGCCTGTGGGACCAAGCCTGAAGCCCTGCGTGGG----	
Vall120-11e201	(321)	CAGCCTGTGGGACCAAGCCTGAAGCCCTGCGTGGG----	
Vall120-Thr202	(321)	CAGCCTGTGGGACCAAGCCTGAAGCCCTGCGTGGG----	
Lys121-Val200	(321)	CAGCCTGTGGGACCAAGCCTGAAGCCCTGCGTGAAGG--	
Consensus	(321)	CAGCCTGTGGGACCAAGCCTGAAGCCCTGCGTGGG	361 400
Leu122-Ser199	(361)	-----GGCAA-----CAGCG	
Vall127-Asn195	(361)	ACCCCCCTGTGCGTGGGGGACGGGAACCTGCAACACCAAGG	
Vall120-11e201B	(357)	-----CG	
Vall120-Ala204	(357)	-----CG	
Vall120-11e201	(357)	-----CG	
Vall120-Thr202	(357)	-----CG	
Lys121-Val200	(359)	-----C-----CCCCG	
Consensus	(361)	CG	401 440
Leu122-Ser199	(371)	TGATCACCACAGGCTGCCCAAGGTGAGCTTCGAGCCCAT	
Vall127-Asn195	(401)	TGATCACCACAGGCTGCCCAAGGTGAGCTTCGAGCCCAT	
Vall120-11e201B	(359)	GCATCACCACAGGCTGCCCAAGGTGAGCTTCGAGCCCAT	
Vall120-Ala204	(357)	-----CGCCGGGGCTGCCCAAGGTGAGCTTCGAGCCCAT	
Vall120-11e201	(359)	GCATCACCACAGGCTGCCCAAGGTGAGCTTCGAGCCCAT	
Vall120-Thr202	(359)	GGCCACCCAGGCTGCCCAAGGTGAGCTTCGAGCCCAT	
Lys121-Val200	(365)	TGATCACCACAGGCTGCCCAAGGTGAGCTTCGAGCCCAT	
Consensus	(401)	ATCACCACAGGCTGCCCAAGGTGAGCTTCGAGCCCAT	441 480
Leu122-Ser199	(411)	CCCCATCCACTACTGCGCCCGCGCGGCTTCGCCATCTCTG	
Vall127-Asn195	(441)	CCCCATCCACTACTGCGCCCGCGCGGCTTCGCCATCTCTG	
Vall120-11e201B	(399)	CCCCATCCACTACTGCGCCCGCGCGGCTTCGCCATCTCTG	
Vall120-Ala204	(393)	CCCCATCCACTACTGCGCCCGCGCGGCTTCGCCATCTCTG	
Vall120-11e201	(399)	CCCCATCCACTACTGCGCCCGCGCGGCTTCGCCATCTCTG	
Vall120-Thr202	(399)	CCCCATCCACTACTGCGCCCGCGCGGCTTCGCCATCTCTG	
Lys121-Val200	(405)	CCCCATCCACTACTGCGCCCGCGCGGCTTCGCCATCTCTG	
Consensus	(441)	CCCCATCCACTACTGCGCCCGCGCGGCTTCGCCATCTCTG	481 520
Leu122-Ser199	(451)	AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTCGCA	
Vall127-Asn195	(481)	AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTCGCA	
Vall120-11e201B	(439)	AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTCGCA	
Vall120-Ala204	(433)	AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTCGCA	
Vall120-11e201	(439)	AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTCGCA	

FIG. 3B

Val120-Thr202	(439)	AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCA	
Lys121-Val200	(445)	AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCA	
Consensus	(481)	AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCA	560
Leu122-Ser199	(491)	CCAACGTGAGCACCCTGCAGTGCACCCACGGCATCCGCC	
Val1127-Asn195	(521)	CCAACGTGAGCACCCTGCAGTGCACCCACGGCATCCGCC	
Val1120-Ile201B	(479)	CCAACGTGAGCACCCTGCAGTGCACCCACGGCATCCGCC	
Val1120-Ala204	(473)	CCAACGTGAGCACCCTGCAGTGCACCCACGGCATCCGCC	
Val1120-Ile201	(479)	CCAACGTGAGCACCCTGCAGTGCACCCACGGCATCCGCC	
Val1120-Thr202	(479)	CCAACGTGAGCACCCTGCAGTGCACCCACGGCATCCGCC	
Lys121-Val200	(485)	CCAACGTGAGCACCCTGCAGTGCACCCACGGCATCCGCC	
Consensus	(521)	CCAACGTGAGCACCCTGCAGTGCACCCACGGCATCCGCC	600
Leu122-Ser199	(531)	CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC	
Val1127-Asn195	(561)	CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC	
Val1120-Ile201B	(519)	CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC	
Val1120-Ala204	(513)	CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC	
Val1120-Ile201	(519)	CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC	
Val1120-Thr202	(519)	CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC	
Lys121-Val200	(525)	CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC	
Consensus	(561)	CGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCC	640
Leu122-Ser199	(571)	GAGGAGGGGCGTGGTGATCCGCAAGCGAGAACTTCACCGACA	
Val1127-Asn195	(601)	GAGGAGGGGCGTGGTGATCCGCAAGCGAGAACTTCACCGACA	
Val1120-Ile201B	(559)	GAGGAGGGGCGTGGTGATCCGCAAGCGAGAACTTCACCGACA	
Val1120-Ala204	(553)	GAGGAGGGGCGTGGTGATCCGCAAGCGAGAACTTCACCGACA	
Val1120-Ile201	(559)	GAGGAGGGGCGTGGTGATCCGCAAGCGAGAACTTCACCGACA	
Val1120-Thr202	(559)	GAGGAGGGGCGTGGTGATCCGCAAGCGAGAACTTCACCGACA	
Lys121-Val200	(565)	GAGGAGGGGCGTGGTGATCCGCAAGCGAGAACTTCACCGACA	
Consensus	(601)	GAGGAGGGGCGTGGTGATCCGCAAGCGAGAACTTCACCGACA	680
Leu122-Ser199	(611)	ACGCCAAGACCATCATCGTGACGCTGAAGGAGAGCGGTGGA	
Val1127-Asn195	(641)	ACGCCAAGACCATCATCGTGACGCTGAAGGAGAGCGGTGGA	
Val1120-Ile201B	(599)	ACGCCAAGACCATCATCGTGACGCTGAAGGAGAGCGGTGGA	
Val1120-Ala204	(593)	ACGCCAAGACCATCATCGTGACGCTGAAGGAGAGCGGTGGA	
Val1120-Ile201	(599)	ACGCCAAGACCATCATCGTGACGCTGAAGGAGAGCGGTGGA	
Val1120-Thr202	(599)	ACGCCAAGACCATCATCGTGACGCTGAAGGAGAGCGGTGGA	
Lys121-Val200	(605)	ACGCCAAGACCATCATCGTGACGCTGAAGGAGAGCGGTGGA	
Consensus	(641)	ACGCCAAGACCATCATCGTGACGCTGAAGGAGAGCGGTGGA	720
Leu122-Ser199	(651)	GATCAACTGCACCGCCGCCCAACAAACACCCGCAAGAGC	
Val1127-Asn195	(681)	GATCAACTGCACCGCCGCCCAACAAACACCCGCAAGAGC	
Val1120-Ile201B	(639)	GATCAACTGCACCGCCGCCCAACAAACACCCGCAAGAGC	
Val1120-Ala204	(633)	GATCAACTGCACCGCCGCCCAACAAACACCCGCAAGAGC	
Val1120-Ile201	(639)	GATCAACTGCACCGCCGCCCAACAAACACCCGCAAGAGC	
Val1120-Thr202	(639)	GATCAACTGCACCGCCGCCCAACAAACACCCGCAAGAGC	
Lys121-Val200	(645)	GATCAACTGCACCGCCGCCCAACAAACACCCGCAAGAGC	
Consensus	(681)	GATCAACTGCACCGCCGCCCAACAAACACCCGCAAGAGC	760
Leu122-Ser199	(691)	ATCACCATCGGCCCGCGCGCGCTTCTACGCCACCGGCG	
Val1127-Asn195	(721)	ATCACCATCGGCCCGCGCGCGCTTCTACGCCACCGGCG	
Val1120-Ile201B	(679)	ATCACCATCGGCCCGCGCGCGCTTCTACGCCACCGGCG	
Val1120-Ala204	(673)	ATCACCATCGGCCCGCGCGCGCTTCTACGCCACCGGCG	
Val1120-Ile201	(679)	ATCACCATCGGCCCGCGCGCGCTTCTACGCCACCGGCG	
Val1120-Thr202	(679)	ATCACCATCGGCCCGCGCGCGCTTCTACGCCACCGGCG	
Lys121-Val200	(685)	ATCACCATCGGCCCGCGCGCGCTTCTACGCCACCGGCG	
Consensus	(721)	ATCACCATCGGCCCGCGCGCGCTTCTACGCCACCGGCG	

FIG. 3C

	761	800
Leu122-Ser199	(731) ACATCATCGGGGACATCCGGCAGGCCACTGCAACATCAG	
Val1127-Asn195	(761) ACATCATCGGGGACATCCGGCAGGCCACTGCAACATCAG	
Val1120-11e201B	(719) ACATCATCGGGGACATCCGGCAGGCCACTGCAACATCAG	
Val1120-Ala204	(713) ACATCATCGGGGACATCCGGCAGGCCACTGCAACATCAG	
Val1120-11e201	(719) ACATCATCGGGGACATCCGGCAGGCCACTGCAACATCAG	
Val1120-Thr202	(719) ACATCATCGGGGACATCCGGCAGGCCACTGCAACATCAG	
Lys121-Val1200	(725) ACATCATCGGGGACATCCGGCAGGCCACTGCAACATCAG	
Consensus	(761) ACATCATCGGGGACATCCGGCAGGCCACTGCAACATCAG	840
Leu122-Ser199	(771) CGGCGAGAAGTGGAAACAACACCCCTGAAGCAGATCGTGACC	
Val1127-Asn195	(801) CGGCGAGAAGTGGAAACAACACCCCTGAAGCAGATCGTGACC	
Val1120-11e201B	(759) CGGCGAGAAGTGGAAACAACACCCCTGAAGCAGATCGTGACC	
Val1120-Ala204	(753) CGGCGAGAAGTGGAAACAACACCCCTGAAGCAGATCGTGACC	
Val1120-11e201	(759) CGGCGAGAAGTGGAAACAACACCCCTGAAGCAGATCGTGACC	
Val1120-Thr202	(759) CGGCGAGAAGTGGAAACAACACCCCTGAAGCAGATCGTGACC	
Lys121-Val1200	(765) CGGCGAGAAGTGGAAACAACACCCCTGAAGCAGATCGTGACC	
Consensus	(801) CGGCGAGAAGTGGAAACAACACCCCTGAAGCAGATCGTGACC	880
Leu122-Ser199	(811) AAGCTGCAGGCCAGTTTCGGCAACAAGACCATCGTGTTC	
Val1127-Asn195	(841) AAGCTGCAGGCCAGTTTCGGCAACAAGACCATCGTGTTC	
Val1120-11e201B	(799) AAGCTGCAGGCCAGTTTCGGCAACAAGACCATCGTGTTC	
Val1120-Ala204	(793) AAGCTGCAGGCCAGTTTCGGCAACAAGACCATCGTGTTC	
Val1120-11e201	(799) AAGCTGCAGGCCAGTTTCGGCAACAAGACCATCGTGTTC	
Val1120-Thr202	(799) AAGCTGCAGGCCAGTTTCGGCAACAAGACCATCGTGTTC	
Lys121-Val1200	(805) AAGCTGCAGGCCAGTTTCGGCAACAAGACCATCGTGTTC	
Consensus	(841) AAGCTGCAGGCCAGTTTCGGCAACAAGACCATCGTGTTC	920
Leu122-Ser199	(851) AGCAGAGCAGCGGGCGGCGACCCCGAGATCGTGATGCACAG	
Val1127-Asn195	(881) AGCAGAGCAGCGGGCGGCGACCCCGAGATCGTGATGCACAG	
Val1120-11e201B	(839) AGCAGAGCAGCGGGCGGCGACCCCGAGATCGTGATGCACAG	
Val1120-Ala204	(833) AGCAGAGCAGCGGGCGGCGACCCCGAGATCGTGATGCACAG	
Val1120-11e201	(839) AGCAGAGCAGCGGGCGGCGACCCCGAGATCGTGATGCACAG	
Val1120-Thr202	(839) AGCAGAGCAGCGGGCGGCGACCCCGAGATCGTGATGCACAG	
Lys121-Val1200	(845) AGCAGAGCAGCGGGCGGCGACCCCGAGATCGTGATGCACAG	
Consensus	(881) AGCAGAGCAGCGGGCGGCGACCCCGAGATCGTGATGCACAG	960
Leu122-Ser199	(891) CTTCAACTGCGGGCGGCGAGTTCTTCTACTGCAACAGCACC	
Val1127-Asn195	(921) CTTCAACTGCGGGCGGCGAGTTCTTCTACTGCAACAGCACC	
Val1120-11e201B	(879) CTTCAACTGCGGGCGGCGAGTTCTTCTACTGCAACAGCACC	
Val1120-Ala204	(873) CTTCAACTGCGGGCGGCGAGTTCTTCTACTGCAACAGCACC	
Val1120-11e201	(879) CTTCAACTGCGGGCGGCGAGTTCTTCTACTGCAACAGCACC	
Val1120-Thr202	(879) CTTCAACTGCGGGCGGCGAGTTCTTCTACTGCAACAGCACC	
Lys121-Val1200	(885) CTTCAACTGCGGGCGGCGAGTTCTTCTACTGCAACAGCACC	
Consensus	(921) CTTCAACTGCGGGCGGCGAGTTCTTCTACTGCAACAGCACC	1000
Leu122-Ser199	(931) CAGCTGTTCAACAGCACCTGGAAACAACACCATCGGCCCCA	
Val1127-Asn195	(961) CAGCTGTTCAACAGCACCTGGAAACAACACCATCGGCCCCA	
Val1120-11e201B	(919) CAGCTGTTCAACAGCACCTGGAAACAACACCATCGGCCCCA	
Val1120-Ala204	(913) CAGCTGTTCAACAGCACCTGGAAACAACACCATCGGCCCCA	
Val1120-11e201	(919) CAGCTGTTCAACAGCACCTGGAAACAACACCATCGGCCCCA	
Val1120-Thr202	(919) CAGCTGTTCAACAGCACCTGGAAACAACACCATCGGCCCCA	
Lys121-Val1200	(925) CAGCTGTTCAACAGCACCTGGAAACAACACCATCGGCCCCA	
Consensus	(961) CAGCTGTTCAACAGCACCTGGAAACAACACCATCGGCCCCA	1040
Leu122-Ser199	(971) ACAACACCAACGGCACCATCAACCTGCCCTGCCGATCAA	
Val1127-Asn195	(1001) ACAACACCAACGGCACCATCAACCTGCCCTGCCGATCAA	

FIG. 3D

Val120-11e201B	(959)	ACAACACCAACGGGACCATCACCTGCCCTGCCGCATCAA
Val120-Ala204	(953)	ACAACACCAACGGGACCATCACCTGCCCTGCCGCATCAA
Val120-11e201	(959)	ACAACACCAACGGGACCATCACCTGCCCTGCCGCATCAA
Val120-Thr202	(959)	ACAACACCAACGGGACCATCACCTGCCCTGCCGCATCAA
Lys121-Val200	(965)	ACAACACCAACGGGACCATCACCTGCCCTGCCGCATCAA
Consensus	(1001)	ACAACACCAACGGGACCATCACCTGCCCTGCCGCATCAA
Leu122-Ser199	(1011)	GCAGATCATCAACCGCTGGCAGGAGTGGGCAAGGCCATG
Val127-Asn195	(1041)	GCAGATCATCAACCGCTGGCAGGAGTGGGCAAGGCCATG
Val120-11e201B	(999)	GCAGATCATCAACCGCTGGCAGGAGTGGGCAAGGCCATG
Val120-Ala204	(993)	GCAGATCATCAACCGCTGGCAGGAGTGGGCAAGGCCATG
Val120-11e201	(999)	GCAGATCATCAACCGCTGGCAGGAGTGGGCAAGGCCATG
Val120-Thr202	(999)	GCAGATCATCAACCGCTGGCAGGAGTGGGCAAGGCCATG
Lys121-Val200	(1005)	GCAGATCATCAACCGCTGGCAGGAGTGGGCAAGGCCATG
Consensus	(1041)	GCAGATCATCAACCGCTGGCAGGAGTGGGCAAGGCCATG
Leu122-Ser199	(1051)	TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val127-Asn195	(1081)	TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val120-11e201B	(1039)	TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val120-Ala204	(1033)	TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val120-11e201	(1039)	TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Val120-Thr202	(1039)	TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Lys121-Val200	(1045)	TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Consensus	(1081)	TACGCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCA
Leu122-Ser199	(1091)	ACATCACCGGCTGCTGCTGACCCGCGACGGCGGCAAGGA
Val127-Asn195	(1121)	ACATCACCGGCTGCTGCTGACCCGCGACGGCGGCAAGGA
Val120-11e201B	(1079)	ACATCACCGGCTGCTGCTGACCCGCGACGGCGGCAAGGA
Val120-Ala204	(1073)	ACATCACCGGCTGCTGCTGACCCGCGACGGCGGCAAGGA
Val120-11e201	(1079)	ACATCACCGGCTGCTGCTGACCCGCGACGGCGGCAAGGA
Val120-Thr202	(1079)	ACATCACCGGCTGCTGCTGACCCGCGACGGCGGCAAGGA
Lys121-Val200	(1085)	ACATCACCGGCTGCTGCTGACCCGCGACGGCGGCAAGGA
Consensus	(1121)	ACATCACCGGCTGCTGCTGACCCGCGACGGCGGCAAGGA
Leu122-Ser199	(1131)	GATCAGCAACACCAACCGAGATCTTCCGCCCGGGGCGCGGC
Val127-Asn195	(1161)	GATCAGCAACACCAACCGAGATCTTCCGCCCGGGGCGCGGC
Val120-11e201B	(1119)	GATCAGCAACACCAACCGAGATCTTCCGCCCGGGGCGCGGC
Val120-Ala204	(1113)	GATCAGCAACACCAACCGAGATCTTCCGCCCGGGGCGCGGC
Val120-11e201	(1119)	GATCAGCAACACCAACCGAGATCTTCCGCCCGGGGCGCGGC
Val120-Thr202	(1119)	GATCAGCAACACCAACCGAGATCTTCCGCCCGGGGCGCGGC
Lys121-Val200	(1125)	GATCAGCAACACCAACCGAGATCTTCCGCCCGGGGCGCGGC
Consensus	(1161)	GATCAGCAACACCAACCGAGATCTTCCGCCCGGGGCGCGGC
Leu122-Ser199	(1171)	GACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACA
Val127-Asn195	(1201)	GACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACA
Val120-11e201B	(1159)	GACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACA
Val120-Ala204	(1153)	GACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACA
Val120-11e201	(1159)	GACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACA
Val120-Thr202	(1159)	GACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACA
Lys121-Val200	(1165)	GACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACA
Consensus	(1201)	GACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACA
Leu122-Ser199	(1211)	AGGTGGTGAAGATCGAGCCCTGGGCGTGGCCCCCACCAA
Val127-Asn195	(1241)	AGGTGGTGAAGATCGAGCCCTGGGCGTGGCCCCCACCAA
Val120-11e201B	(1199)	AGGTGGTGAAGATCGAGCCCTGGGCGTGGCCCCCACCAA
Val120-Ala204	(1193)	AGGTGGTGAAGATCGAGCCCTGGGCGTGGCCCCCACCAA
Val120-11e201	(1199)	AGGTGGTGAAGATCGAGCCCTGGGCGTGGCCCCCACCAA

FIG. 3E



Val120-Thr202	(1199)	AGGTGGTGAAGATCGAGCCCCCTGGGCGTG6CCCCACCAA	
Lys121-Val200	(1205)	AGGTGGTGAAGATCGAGCCCCCTGGGCGTG6CCCCACCAA	
Consensus	(1241)	AGGTGGTGAAGATCGAGCCCCCTGGGCGTG6CCCCACCAA	1281 1320
Leu122-Ser199	(1251)	GGCCAAAGCGCCGCTGGTGACGCGCAGAAAGCGCCCGTG	
Val127-Asn195	(1281)	GGCCAAAGCGCCGCTGGTGACGCGCAGAAAGCGCCCGTG	
Val120-Ile201B	(1239)	GGCCAAAGCGCCGCTGGTGACGCGCAGAAAGCGCCCGTG	
Val120-Ala204	(1233)	GGCCAAAGCGCCGCTGGTGACGCGCAGAAAGCGCCCGTG	
Val120-Ile201	(1239)	GGCCAAAGCGCCGCTGGTGACGCGCAGAAAGCGCCCGTG	
Val120-Thr202	(1239)	GGCCAAAGCGCCGCTGGTGACGCGCAGAAAGCGCCCGTG	
Lys121-Val200	(1245)	GGCCAAAGCGCCGCTGGTGACGCGCAGAAAGCGCCCGTG	
Consensus	(1281)	GGCCAAAGCGCCGCTGGTGACGCGCAGAAAGCGCCCGTG	1321 1360
Leu122-Ser199	(1291)	ACCTGGGGGCCATGTTCTTGGGCTTCTGGGCGCCGCGG	
Val127-Asn195	(1321)	ACCTGGGGGCCATGTTCTTGGGCTTCTGGGCGCCGCGG	
Val120-Ile201B	(1279)	ACCTGGGGGCCATGTTCTTGGGCTTCTGGGCGCCGCGG	
Val120-Ala204	(1273)	ACCTGGGGGCCATGTTCTTGGGCTTCTGGGCGCCGCGG	
Val120-Ile201	(1279)	ACCTGGGGGCCATGTTCTTGGGCTTCTGGGCGCCGCGG	
Val120-Thr202	(1279)	ACCTGGGGGCCATGTTCTTGGGCTTCTGGGCGCCGCGG	
Lys121-Val200	(1285)	ACCTGGGGGCCATGTTCTTGGGCTTCTGGGCGCCGCGG	
Consensus	(1321)	ACCTGGGGGCCATGTTCTTGGGCTTCTGGGCGCCGCGG	1361 1400
Leu122-Ser199	(1331)	GCAGCACCATGGGCGCCGCGAGCCTGACCTGACCGTGCA	
Val127-Asn195	(1361)	GCAGCACCATGGGCGCCGCGAGCCTGACCTGACCGTGCA	
Val120-Ile201B	(1319)	GCAGCACCATGGGCGCCGCGAGCCTGACCTGACCGTGCA	
Val120-Ala204	(1313)	GCAGCACCATGGGCGCCGCGAGCCTGACCTGACCGTGCA	
Val120-Ile201	(1319)	GCAGCACCATGGGCGCCGCGAGCCTGACCTGACCGTGCA	
Val120-Thr202	(1319)	GCAGCACCATGGGCGCCGCGAGCCTGACCTGACCGTGCA	
Lys121-Val200	(1325)	GCAGCACCATGGGCGCCGCGAGCCTGACCTGACCGTGCA	
Consensus	(1361)	GCAGCACCATGGGCGCCGCGAGCCTGACCTGACCGTGCA	1401 1440
Leu122-Ser199	(1371)	GGCCCGCCAGCTGCTGAGCGGCATCGTGACGACGAGAAC	
Val127-Asn195	(1401)	GGCCCGCCAGCTGCTGAGCGGCATCGTGACGACGAGAAC	
Val120-Ile201B	(1359)	GGCCCGCCAGCTGCTGAGCGGCATCGTGACGACGAGAAC	
Val120-Ala204	(1353)	GGCCCGCCAGCTGCTGAGCGGCATCGTGACGACGAGAAC	
Val120-Ile201	(1359)	GGCCCGCCAGCTGCTGAGCGGCATCGTGACGACGAGAAC	
Val120-Thr202	(1359)	GGCCCGCCAGCTGCTGAGCGGCATCGTGACGACGAGAAC	
Lys121-Val200	(1365)	GGCCCGCCAGCTGCTGAGCGGCATCGTGACGACGAGAAC	
Consensus	(1401)	GGCCCGCCAGCTGCTGAGCGGCATCGTGACGACGAGAAC	1441 1480
Leu122-Ser199	(1411)	AACCTGCTGCGCGCCATCGAGGCCAAGCAGCAACCTGCTGC	
Val127-Asn195	(1441)	AACCTGCTGCGCGCCATCGAGGCCAAGCAGCAACCTGCTGC	
Val120-Ile201B	(1399)	AACCTGCTGCGCGCCATCGAGGCCAAGCAGCAACCTGCTGC	
Val120-Ala204	(1393)	AACCTGCTGCGCGCCATCGAGGCCAAGCAGCAACCTGCTGC	
Val120-Ile201	(1399)	AACCTGCTGCGCGCCATCGAGGCCAAGCAGCAACCTGCTGC	
Val120-Thr202	(1399)	AACCTGCTGCGCGCCATCGAGGCCAAGCAGCAACCTGCTGC	
Lys121-Val200	(1405)	AACCTGCTGCGCGCCATCGAGGCCAAGCAGCAACCTGCTGC	
Consensus	(1441)	AACCTGCTGCGCGCCATCGAGGCCAAGCAGCAACCTGCTGC	1481 1520
Leu122-Ser199	(1451)	AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGT	
Val127-Asn195	(1481)	AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGT	
Val120-Ile201B	(1439)	AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGT	
Val120-Ala204	(1433)	AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGT	
Val120-Ile201	(1439)	AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGT	
Val120-Thr202	(1439)	AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGT	
Lys121-Val200	(1445)	AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGT	
Consensus	(1481)	AGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGT	

		1521	1560
Leu122-Ser199	(1491)	GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG	
Val1127-Asn195	(1521)	GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG	
Val1120-11e201B	(1479)	GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG	
Val1120-Ala204	(1473)	GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG	
Val1120-11e201	(1479)	GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG	
Val1120-Thr202	(1479)	GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG	
Lys121-Val200	(1485)	GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG	
Consensus	(1521)	GCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTG	
		1561	1600
Leu122-Ser199	(1531)	GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG	
Val1127-Asn195	(1561)	GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG	
Val1120-11e201B	(1519)	GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG	
Val1120-Ala204	(1513)	GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG	
Val1120-11e201	(1519)	GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG	
Val1120-Thr202	(1519)	GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG	
Lys121-Val200	(1525)	GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG	
Consensus	(1561)	GGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCG	
		1601	1640
Leu122-Ser199	(1571)	CGGTGCCCTGGAAACGCCAGCTGGAGCAACAGAGCGCTTGA	
Val1127-Asn195	(1601)	CGGTGCCCTGGAAACGCCAGCTGGAGCAACAGAGCGCTTGA	
Val1120-11e201B	(1559)	CGGTGCCCTGGAAACGCCAGCTGGAGCAACAGAGCGCTTGA	
Val1120-Ala204	(1553)	CGGTGCCCTGGAAACGCCAGCTGGAGCAACAGAGCGCTTGA	
Val1120-11e201	(1559)	CGGTGCCCTGGAAACGCCAGCTGGAGCAACAGAGCGCTTGA	
Val1120-Thr202	(1559)	CGGTGCCCTGGAAACGCCAGCTGGAGCAACAGAGCGCTTGA	
Lys121-Val200	(1565)	CGGTGCCCTGGAAACGCCAGCTGGAGCAACAGAGCGCTTGA	
Consensus	(1601)	CGGTGCCCTGGAAACGCCAGCTGGAGCAACAGAGCGCTTGA	
		1641	1680
Leu122-Ser199	(1611)	CCAGATCTGGAAACACATGACCTGGATGGAGTGGGAGCGGC	
Val1127-Asn195	(1641)	CCAGATCTGGAAACACATGACCTGGATGGAGTGGGAGCGGC	
Val1120-11e201B	(1599)	CCAGATCTGGAAACACATGACCTGGATGGAGTGGGAGCGGC	
Val1120-Ala204	(1593)	CCAGATCTGGAAACACATGACCTGGATGGAGTGGGAGCGGC	
Val1120-11e201	(1599)	CCAGATCTGGAAACACATGACCTGGATGGAGTGGGAGCGGC	
Val1120-Thr202	(1599)	CCAGATCTGGAAACACATGACCTGGATGGAGTGGGAGCGGC	
Lys121-Val200	(1605)	CCAGATCTGGAAACACATGACCTGGATGGAGTGGGAGCGGC	
Consensus	(1641)	CCAGATCTGGAAACACATGACCTGGATGGAGTGGGAGCGGC	
		1681	1720
Leu122-Ser199	(1651)	GAGATCGACAACTACACCAACCTGATCTACACCCGTGATCG	
Val1127-Asn195	(1681)	GAGATCGACAACTACACCAACCTGATCTACACCCGTGATCG	
Val1120-11e201B	(1639)	GAGATCGACAACTACACCAACCTGATCTACACCCGTGATCG	
Val1120-Ala204	(1633)	GAGATCGACAACTACACCAACCTGATCTACACCCGTGATCG	
Val1120-11e201	(1639)	GAGATCGACAACTACACCAACCTGATCTACACCCGTGATCG	
Val1120-Thr202	(1639)	GAGATCGACAACTACACCAACCTGATCTACACCCGTGATCG	
Lys121-Val200	(1645)	GAGATCGACAACTACACCAACCTGATCTACACCCGTGATCG	
Consensus	(1681)	GAGATCGACAACTACACCAACCTGATCTACACCCGTGATCG	
		1721	1760
Leu122-Ser199	(1691)	AGGAGAGCCAGAACGACGAGGAGAGAACGAGCAGGAGCT	
Val1127-Asn195	(1721)	AGGAGAGCCAGAACGACGAGGAGAGAACGAGCAGGAGCT	
Val1120-11e201B	(1679)	AGGAGAGCCAGAACGACGAGGAGAGAACGAGCAGGAGCT	
Val1120-Ala204	(1673)	AGGAGAGCCAGAACGACGAGGAGAGAACGAGCAGGAGCT	
Val1120-11e201	(1679)	AGGAGAGCCAGAACGACGAGGAGAGAACGAGCAGGAGCT	
Val1120-Thr202	(1679)	AGGAGAGCCAGAACGACGAGGAGAGAACGAGCAGGAGCT	
Lys121-Val200	(1685)	AGGAGAGCCAGAACGACGAGGAGAGAACGAGCAGGAGCT	
Consensus	(1721)	AGGAGAGCCAGAACGACGAGGAGAGAACGAGCAGGAGCT	
		1761	1800
Leu122-Ser199	(1731)	GCTGGAGCTGGACAAGTGGGCGACCTGTGGAACCTGGTTC	
Val1127-Asn195	(1761)	GCTGGAGCTGGACAAGTGGGCGACCTGTGGAACCTGGTTC	

FIG. 3G

Val120-11e201B	(1719)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTTC	
Val120-Ala204	(1713)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTTC	
Val120-11e201	(1719)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTTC	
Val120-Thr202	(1719)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTTC	
Lys121-Val200	(1725)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTTC	
Consensus	(1761)	GCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTTC	1801 1840
Leu122-Ser199	(1771)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTTCATCA	
Val127-Asn195	(1801)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTTCATCA	
Val120-11e201B	(1759)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTTCATCA	
Val120-Ala204	(1753)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTTCATCA	
Val120-11e201	(1759)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTTCATCA	
Val120-Thr202	(1759)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTTCATCA	
Lys121-Val200	(1765)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTTCATCA	
Consensus	(1801)	GACATCAGCAAGTGGCTGTGGTACATCAAGATCTTTCATCA	1841 1880
Leu122-Ser199	(1811)	TGATCGTGGGCGGCCCTGGTGGGCCCTGCGCATCGTGTTCAC	
Val127-Asn195	(1841)	TGATCGTGGGCGGCCCTGGTGGGCCCTGCGCATCGTGTTCAC	
Val120-11e201B	(1799)	TGATCGTGGGCGGCCCTGGTGGGCCCTGCGCATCGTGTTCAC	
Val120-Ala204	(1793)	TGATCGTGGGCGGCCCTGGTGGGCCCTGCGCATCGTGTTCAC	
Val120-11e201	(1799)	TGATCGTGGGCGGCCCTGGTGGGCCCTGCGCATCGTGTTCAC	
Val120-Thr202	(1799)	TGATCGTGGGCGGCCCTGGTGGGCCCTGCGCATCGTGTTCAC	
Lys121-Val200	(1805)	TGATCGTGGGCGGCCCTGGTGGGCCCTGCGCATCGTGTTCAC	
Consensus	(1841)	TGATCGTGGGCGGCCCTGGTGGGCCCTGCGCATCGTGTTCAC	1881 1920
Leu122-Ser199	(1851)	CGTCTGAGCATCGTGAACCGCGCTGCGCCAGGGGCTACAGC	
Val127-Asn195	(1881)	CGTCTGAGCATCGTGAACCGCGCTGCGCCAGGGGCTACAGC	
Val120-11e201B	(1839)	CGTCTGAGCATCGTGAACCGCGCTGCGCCAGGGGCTACAGC	
Val120-Ala204	(1833)	CGTCTGAGCATCGTGAACCGCGCTGCGCCAGGGGCTACAGC	
Val120-11e201	(1839)	CGTCTGAGCATCGTGAACCGCGCTGCGCCAGGGGCTACAGC	
Val120-Thr202	(1839)	CGTCTGAGCATCGTGAACCGCGCTGCGCCAGGGGCTACAGC	
Lys121-Val200	(1845)	CGTCTGAGCATCGTGAACCGCGCTGCGCCAGGGGCTACAGC	
Consensus	(1881)	CGTCTGAGCATCGTGAACCGCGCTGCGCCAGGGGCTACAGC	1921 1960
Leu122-Ser199	(1891)	CCCTTGAGCTTCCAGACCCCGCTTCCCGCCCCCGCGGCC	
Val127-Asn195	(1921)	CCCTTGAGCTTCCAGACCCCGCTTCCCGCCCCCGCGGCC	
Val120-11e201B	(1879)	CCCTTGAGCTTCCAGACCCCGCTTCCCGCCCCCGCGGCC	
Val120-Ala204	(1873)	CCCTTGAGCTTCCAGACCCCGCTTCCCGCCCCCGCGGCC	
Val120-11e201	(1879)	CCCTTGAGCTTCCAGACCCCGCTTCCCGCCCCCGCGGCC	
Val120-Thr202	(1879)	CCCTTGAGCTTCCAGACCCCGCTTCCCGCCCCCGCGGCC	
Lys121-Val200	(1885)	CCCTTGAGCTTCCAGACCCCGCTTCCCGCCCCCGCGGCC	
Consensus	(1921)	CCCTTGAGCTTCCAGACCCCGCTTCCCGCCCCCGCGGCC	1961 2000
Leu122-Ser199	(1931)	CCGACCGCCCGGAGGGCATCGAGGAGGAGGGCGCGGAGCG	
Val127-Asn195	(1961)	CCGACCGCCCGGAGGGCATCGAGGAGGAGGGCGCGGAGCG	
Val120-11e201B	(1919)	CCGACCGCCCGGAGGGCATCGAGGAGGAGGGCGCGGAGCG	
Val120-Ala204	(1913)	CCGACCGCCCGGAGGGCATCGAGGAGGAGGGCGCGGAGCG	
Val120-11e201	(1919)	CCGACCGCCCGGAGGGCATCGAGGAGGAGGGCGCGGAGCG	
Val120-Thr202	(1919)	CCGACCGCCCGGAGGGCATCGAGGAGGAGGGCGCGGAGCG	
Lys121-Val200	(1925)	CCGACCGCCCGGAGGGCATCGAGGAGGAGGGCGCGGAGCG	
Consensus	(1961)	CCGACCGCCCGGAGGGCATCGAGGAGGAGGGCGCGGAGCG	2001 2040
Leu122-Ser199	(1971)	CGACCGGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG	
Val127-Asn195	(2001)	CGACCGGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG	
Val120-11e201B	(1959)	CGACCGGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG	
Val120-Ala204	(1953)	CGACCGGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG	
Val120-11e201	(1959)	CGACCGGACCGCAGCAGCCCCCTGGTGCACGGCCTGCTG	

FIG. 3H

Vall120-Thr202	(1959)	CGACCGCGACCGCAGCAGCCCTGGTGACGGCCGTGCTG
Lys121-Val200	(1965)	CGACCGCGACCGCAGCAGCCCTGGTGACGGCCGTGCTG
Consensus	(2001)	CGACCGCGACCGCAGCAGCCCTGGTGACGGCCGTGCTG
		2041 2080
Leu122-Ser199	(2011)	GCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Val127-Asn195	(2041)	GCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Vall120-Ile201B	(1999)	GCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Val120-Ala204	(1993)	GCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Val120-Ile201	(1999)	GCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Val120-Thr202	(1999)	GCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Lys121-Val200	(2005)	GCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
Consensus	(2041)	GCCTGATCTGGGACGACCTGCGCAGCCTGTGCCTGTTCA
		2081 2120
Leu122-Ser199	(2051)	GCTACACCGCCTGCGGACCTGATCTGATCGCGGCCCG
Val127-Asn195	(2081)	GCTACACCGCCTGCGGACCTGATCTGATCGCGGCCCG
Vall120-Ile201B	(2039)	GCTACACCGCCTGCGGACCTGATCTGATCGCGGCCCG
Val120-Ala204	(2033)	GCTACACCGCCTGCGGACCTGATCTGATCGCGGCCCG
Val120-Ile201	(2039)	GCTACACCGCCTGCGGACCTGATCTGATCGCGGCCCG
Val120-Thr202	(2039)	GCTACACCGCCTGCGGACCTGATCTGATCGCGGCCCG
Lys121-Val200	(2045)	GCTACACCGCCTGCGGACCTGATCTGATCGCGGCCCG
Consensus	(2081)	GCTACACCGCCTGCGGACCTGATCTGATCGCGGCCCG
		2121 2160
Leu122-Ser199	(2091)	CATCTGGAGCTGCTGGGCCCGCGGCTGGGAGGCCCTG
Val127-Asn195	(2121)	CATCTGGAGCTGCTGGGCCCGCGGCTGGGAGGCCCTG
Vall120-Ile201B	(2079)	CATCTGGAGCTGCTGGGCCCGCGGCTGGGAGGCCCTG
Val120-Ala204	(2073)	CATCTGGAGCTGCTGGGCCCGCGGCTGGGAGGCCCTG
Val120-Ile201	(2079)	CATCTGGAGCTGCTGGGCCCGCGGCTGGGAGGCCCTG
Val120-Thr202	(2079)	CATCTGGAGCTGCTGGGCCCGCGGCTGGGAGGCCCTG
Lys121-Val200	(2085)	CATCTGGAGCTGCTGGGCCCGCGGCTGGGAGGCCCTG
Consensus	(2121)	CATCTGGAGCTGCTGGGCCCGCGGCTGGGAGGCCCTG
		2161 2200
Leu122-Ser199	(2131)	AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
Val127-Asn195	(2161)	AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
Vall120-Ile201B	(2119)	AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
Val120-Ala204	(2113)	AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
Val120-Ile201	(2119)	AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
Val120-Thr202	(2119)	AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
Lys121-Val200	(2125)	AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
Consensus	(2161)	AAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGC
		2201 2240
Leu122-Ser199	(2171)	TGAAGAACAGCGCGCTGAGCCTGTTGACGCCATCGCCAT
Val127-Asn195	(2201)	TGAAGAACAGCGCGCTGAGCCTGTTGACGCCATCGCCAT
Vall120-Ile201B	(2159)	TGAAGAACAGCGCGCTGAGCCTGTTGACGCCATCGCCAT
Val120-Ala204	(2153)	TGAAGAACAGCGCGCTGAGCCTGTTGACGCCATCGCCAT
Val120-Ile201	(2159)	TGAAGAACAGCGCGCTGAGCCTGTTGACGCCATCGCCAT
Val120-Thr202	(2159)	TGAAGAACAGCGCGCTGAGCCTGTTGACGCCATCGCCAT
Lys121-Val200	(2165)	TGAAGAACAGCGCGCTGAGCCTGTTGACGCCATCGCCAT
Consensus	(2201)	TGAAGAACAGCGCGCTGAGCCTGTTGACGCCATCGCCAT
		2241 2280
Leu122-Ser199	(2211)	CGCGTGGCCGAGGSCACCGACGATCATCGAGGTGGCC
Val127-Asn195	(2241)	CGCGTGGCCGAGGSCACCGACGATCATCGAGGTGGCC
Vall120-Ile201B	(2199)	CGCGTGGCCGAGGSCACCGACGATCATCGAGGTGGCC
Val120-Ala204	(2193)	CGCGTGGCCGAGGSCACCGACGATCATCGAGGTGGCC
Val120-Ile201	(2199)	CGCGTGGCCGAGGSCACCGACGATCATCGAGGTGGCC
Val120-Thr202	(2199)	CGCGTGGCCGAGGSCACCGACGATCATCGAGGTGGCC
Lys121-Val200	(2205)	CGCGTGGCCGAGGSCACCGACGATCATCGAGGTGGCC
Consensus	(2241)	CGCGTGGCCGAGGSCACCGACGATCATCGAGGTGGCC

		2281		2320
Leu122-Ser199	(2251)	CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA		
Val127-Asn195	(2281)	CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA		
Val120-Ile201B	(2239)	CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA		
Val120-Ala204	(2233)	CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA		
Val120-Ile201	(2239)	CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA		
Val120-Thr202	(2239)	CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA		
Lys121-Val200	(2245)	CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA		
Consensus	(2281)	CAGCGCATCGGCCGCGCCTTCCTGCACATCCCCCGCCGCA		
		2321		2360
Leu122-Ser199	(2291)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAGCG		
Val127-Asn195	(2321)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG--		
Val120-Ile201B	(2279)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAGCG		
Val120-Ala204	(2273)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG--		
Val120-Ile201	(2279)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG--		
Val120-Thr202	(2279)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG--		
Lys121-Val200	(2285)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAGCG		
Consensus	(2321)	TCCGCCAGGGCTTCGAGCGCGCCCTGCTGTAACTCGAG		
		2361		
Leu122-Ser199	(2331)	TGCT		
Val127-Asn195	(2359)	----		
Val120-Ile201B	(2319)	TGCT		
Val120-Ala204	(2311)	----		
Val120-Ile201	(2317)	----		
Val120-Thr202	(2317)	----		
Lys121-Val200	(2325)	TGCT		
Consensus	(2361)			

FIG. 3J

WO 00/39303	15 / 65	PCT/US99/31272
	1	40
Ile424-Ala433	(1)	GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCT
Trp427-Gly431	(1)	41 80
Gln422-Tyr435B	(1)	Ile424-Ala433 (41) GATGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCACG
Arg426-Gly431	(1)	Trp427-Gly431 (41) GATGCTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCACG
Ile423-Met434	(1)	Gln422-Tyr435B (41) GATGCTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCACG
Gln422-Tyr435	(1)	Arg426-Gly431 (41) GATGCTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCACG
Arg426-Lys432	(1)	Ile423-Met434 (41) GATGCTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCACG
Arg426-Gly431B	(1)	Gln422-Tyr435 (41) GATGCTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCACG
Asn425-Lys432	(1)	Arg426-Lys432 (41) GATGCTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCACG
Consensus	(1)	81 120
Ile424-Ala433	(81)	Ile424-Ala433 (81) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Trp427-Gly431	(81)	Trp427-Gly431 (81) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Gln422-Tyr435B	(81)	Gln422-Tyr435B (81) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Arg426-Gly431	(81)	Arg426-Gly431 (81) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Ile423-Met434	(81)	Ile423-Met434 (81) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Gln422-Tyr435	(81)	Gln422-Tyr435 (81) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Arg426-Lys432	(81)	Arg426-Lys432 (81) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Arg426-Gly431B	(81)	Arg426-Gly431B (81) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Asn425-Lys432	(81)	Asn425-Lys432 (81) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Consensus	(81)	121 160
Ile424-Ala433	(121)	Ile424-Ala433 (121) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Trp427-Gly431	(121)	Trp427-Gly431 (121) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Gln422-Tyr435B	(121)	Gln422-Tyr435B (121) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Arg426-Gly431	(121)	Arg426-Gly431 (121) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Ile423-Met434	(121)	Ile423-Met434 (121) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Gln422-Tyr435	(121)	Gln422-Tyr435 (121) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Arg426-Lys432	(121)	Arg426-Lys432 (121) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Arg426-Gly431B	(121)	Arg426-Gly431B (121) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Asn425-Lys432	(121)	Asn425-Lys432 (121) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Consensus	(121)	161 200
Ile424-Ala433	(161)	Ile424-Ala433 (161) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Trp427-Gly431	(161)	Trp427-Gly431 (161) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Gln422-Tyr435B	(161)	Gln422-Tyr435B (161) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Arg426-Gly431	(161)	Arg426-Gly431 (161) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Ile423-Met434	(161)	Ile423-Met434 (161) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Gln422-Tyr435	(161)	Gln422-Tyr435 (161) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Arg426-Lys432	(161)	Arg426-Lys432 (161) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Arg426-Gly431B	(161)	Arg426-Gly431B (161) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Asn425-Lys432	(161)	Asn425-Lys432 (161) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG
Consensus	(161)	201 240
Ile424-Ala433	(201)	Ile424-Ala433 (201) GCGCGTGGAGAAGCTGTGGGTGACCGGTACTACCGCGTG

FIG. 4A

FIG. 4B

WO 00/39303	17 / 65	PCT/US99/31272
Arg426-Gly431	(401)	ACGCCCACCAACACCAAGAGCAGCACTGGAAGGAGATGGA
Ile423-Met434	(401)	441 480
Gln422-Tyr435	(401)	Ile424-Ala433 (441)
Arg426-Lys432	(401)	Trp427-Gly431 (441)
Arg426-Gly431B	(401)	Gln422-Tyr435B (441)
Asn425-Lys432	(401)	Arg426-Gly431 (441)
Consensus	(401)	Ile423-Met434 (441)
		Gln422-Tyr435 (441)
		Arg426-Lys432 (441)
		Arg426-Gly431B (441)
		Asn425-Lys432 (441)
		Consensus (441)
		481 520
Ile424-Ala433	(481)	Trp427-Gly431 (481)
Trp427-Gly431	(481)	Gln422-Tyr435B (481)
Gln422-Tyr435B	(481)	Arg426-Gly431 (481)
Arg426-Gly431	(481)	Ile423-Met434 (481)
Ile423-Met434	(481)	Gln422-Tyr435 (481)
Gln422-Tyr435	(481)	Arg426-Lys432 (481)
Arg426-Lys432	(481)	Arg426-Gly431B (481)
Arg426-Gly431B	(481)	Asn425-Lys432 (481)
Asn425-Lys432	(481)	Consensus (481)
Consensus	(481)	521 560
		Ile424-Ala433 (521)
		Trp427-Gly431 (521)
		Gln422-Tyr435B (521)
		Arg426-Gly431 (521)
		Ile423-Met434 (521)
		Gln422-Tyr435 (521)
		Arg426-Lys432 (521)
		Arg426-Gly431B (521)
		Asn425-Lys432 (521)
		Consensus (521)
		561 600
		Ile424-Ala433 (561)
		Trp427-Gly431 (561)
		Gln422-Tyr435B (561)
		Arg426-Gly431 (561)
		Ile423-Met434 (561)
		Gln422-Tyr435 (561)
		Arg426-Lys432 (561)
		Arg426-Gly431B (561)
		Asn425-Lys432 (561)
		Consensus (561)
		601 640
		Ile424-Ala433 (601)
		Trp427-Gly431 (601)
		Gln422-Tyr435B (601)
		Arg426-Gly431 (601)
		Ile423-Met434 (601)

FIG. 4C



FIG. 4D

FIG. 4F

FIG. 4F

FIG. 4G

Gln422-Tyr435B	(1417)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Arg426-Gly431	(1441)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Ile423-Met434	(1423)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Gln422-Tyr435	(1417)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Arg426-Lys432	(1441)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Arg426-Gly431B	(1441)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Asn425-Lys432	(1435)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Consensus	(1441)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
		1481 1520
Ile424-Ala433	(1469)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Trp427-Gly431	(1481)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Gln422-Tyr435B	(1457)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Arg426-Gly431	(1481)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Ile423-Met434	(1463)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Gln422-Tyr435	(1457)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Arg426-Lys432	(1481)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Arg426-Gly431B	(1481)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Asn425-Lys432	(1475)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Consensus	(1481)	TGCAGCGCGAGAAGCGCGCGTGGCCCCACCAAGCGCGTGG
		1521 1560
Ile424-Ala433	(1509)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Trp427-Gly431	(1521)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Gln422-Tyr435B	(1497)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Arg426-Gly431	(1521)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Ile423-Met434	(1503)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Gln422-Tyr435	(1497)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Arg426-Lys432	(1521)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Arg426-Gly431B	(1521)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Asn425-Lys432	(1515)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Consensus	(1521)	CCTGGGCTTCCTGGGCGCGCGCGGCGAGCACCATGGGCGCC
		1561 1600
Ile424-Ala433	(1549)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Trp427-Gly431	(1561)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Gln422-Tyr435B	(1537)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Arg426-Gly431	(1561)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Ile423-Met434	(1543)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Gln422-Tyr435	(1537)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Arg426-Lys432	(1561)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Arg426-Gly431B	(1561)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Asn425-Lys432	(1555)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Consensus	(1561)	CGCAGCCTGACCTGACCGTGCAGGCCCGCCAGCTGCTGA
		1601 1640
Ile424-Ala433	(1589)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Trp427-Gly431	(1601)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Gln422-Tyr435B	(1577)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Arg426-Gly431	(1601)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Ile423-Met434	(1583)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Gln422-Tyr435	(1577)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Arg426-Lys432	(1601)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Arg426-Gly431B	(1601)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Asn425-Lys432	(1595)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Consensus	(1601)	CGCGCATCGTGAGCAGCAGAACCACTGCTGCGCGGCAT
		1641 1680
Ile424-Ala433	(1629)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Trp427-Gly431	(1641)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Gln422-Tyr435B	(1617)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG
Arg426-Gly431	(1641)	CGGCTGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGG

FIG. 4H



FIG. 4J

Asn425-Lys432	(2035)	GTGGGCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGA	2081	2120
Consensus	(2041)			
Ile424-Ala433	(2069)			
Trp427-Gly431	(2081)			
Gln422-Tyr435B	(2057)			
Arg426-Gly431	(2081)			
Ile423-Met434	(2063)			
Gln422-Tyr435	(2057)			
Arg426-Lys432	(2081)			
Arg426-Gly431B	(2081)			
Asn425-Lys432	(2075)			
Consensus	(2081)	ACCGCGTGGCCAGGGCTACAGCCCCCTGAGCTTCCAGAC	2121	2160
Ile424-Ala433	(2109)			
Trp427-Gly431	(2121)			
Gln422-Tyr435B	(2097)			
Arg426-Gly431	(2121)			
Ile423-Met434	(2103)			
Gln422-Tyr435	(2097)			
Arg426-Lys432	(2121)			
Arg426-Gly431B	(2121)			
Asn425-Lys432	(2115)			
Consensus	(2121)	CCGCTTCCCGCCCCCGCGGCCCGACGCCCGCAGGGCG	2161	2200
Ile424-Ala433	(2149)			
Trp427-Gly431	(2161)			
Gln422-Tyr435B	(2137)			
Arg426-Gly431	(2161)			
Ile423-Met434	(2143)			
Gln422-Tyr435	(2137)			
Arg426-Lys432	(2161)			
Arg426-Gly431B	(2161)			
Asn425-Lys432	(2155)			
Consensus	(2161)	ATCGAGGAGGAGGGCGCGAGCGCGACCGCGACCGCAGCA	2201	2240
Ile424-Ala433	(2189)			
Trp427-Gly431	(2201)			
Gln422-Tyr435B	(2177)			
Arg426-Gly431	(2201)			
Ile423-Met434	(2183)			
Gln422-Tyr435	(2177)			
Arg426-Lys432	(2201)			
Arg426-Gly431B	(2201)			
Asn425-Lys432	(2195)			
Consensus	(2201)	GCCCCCTGGTGACGGCCTGCTGGCCCTGATCTGGGACGA	2241	2280
Ile424-Ala433	(2229)			
Trp427-Gly431	(2241)			
Gln422-Tyr435B	(2217)			
Arg426-Gly431	(2241)			
Ile423-Met434	(2223)			
Gln422-Tyr435	(2217)			
Arg426-Lys432	(2241)			
Arg426-Gly431B	(2241)			
Asn425-Lys432	(2235)			
Consensus	(2241)	CCTGCGCAGCCTGTGCTGTTACGCTACACCGCCTGCGC		

FIG. 4K



Ile424-Ala433	(2269)	2281	2320
Trp427-Gly431	(2281)		
Gln422-Tyr435B	(2257)		
Arg426-Gly431	(2281)		
Ile423-Met434	(2263)		
Gln422-Tyr435	(2257)		
Arg426-Lys432	(2281)		
Arg426-Gly431B	(2281)		
Asn425-Lys432	(2275)		
Consensus	(2281)	GACCTGATCCTGATCGCCGCCGCATCGTGGAGCTGCTGG	2321 2360
Ile424-Ala433	(2309)		
Trp427-Gly431	(2321)		
Gln422-Tyr435B	(2297)		
Arg426-Gly431	(2321)		
Ile423-Met434	(2303)		
Gln422-Tyr435	(2297)		
Arg426-Lys432	(2321)		
Arg426-Gly431B	(2321)		
Asn425-Lys432	(2315)		
Consensus	(2321)	GCCGCCGCGCTGGGAGGCCCTGAAGTACTGGGGCAACCT	2361 2400
Ile424-Ala433	(2349)		
Trp427-Gly431	(2361)		
Gln422-Tyr435B	(2337)		
Arg426-Gly431	(2361)		
Ile423-Met434	(2343)		
Gln422-Tyr435	(2337)		
Arg426-Lys432	(2361)		
Arg426-Gly431B	(2361)		
Asn425-Lys432	(2355)		
Consensus	(2361)	GCTCGAGTACTGGATCCAGGAGCTGAAGAACAGCGCGCTG	2401 2440
Ile424-Ala433	(2389)		
Trp427-Gly431	(2401)		
Gln422-Tyr435B	(2377)		
Arg426-Gly431	(2401)		
Ile423-Met434	(2383)		
Gln422-Tyr435	(2377)		
Arg426-Lys432	(2401)		
Arg426-Gly431B	(2401)		
Asn425-Lys432	(2395)		
Consensus	(2401)	AGCCTGTTGACGCCATCGCCATCGCGTGGCCGAGGGCA	2441 2480
Ile424-Ala433	(2429)		
Trp427-Gly431	(2441)		
Gln422-Tyr435B	(2417)		
Arg426-Gly431	(2441)		
Ile423-Met434	(2423)		
Gln422-Tyr435	(2417)		
Arg426-Lys432	(2441)		
Arg426-Gly431B	(2441)		
Asn425-Lys432	(2435)		
Consensus	(2441)	CGCAGCGCATCATCGAGGTGGCCAGCGCATCGCGCGCGC	2481 2520
Ile424-Ala433	(2469)		

FIG. 4L

Trp427-Gly431	(2481)	CTTCTGACATCCCCCGCCGATCCGCCAGGGCTTCGAG -
Gln422-Tyr435B	(2457)	2521 2541
Arg426-Gly431	(2481)	
Ile423-Met434	(2463)	
Gln422-Tyr435	(2457)	
Arg426-Lys432	(2481)	
Arg426-Gly431B	(2481)	
Asn425-Lys432	(2475)	
Consensus	(2481)	
Ile424-Ala433	(2509)	
Trp427-Gly431	(2521)	
Gln422-Tyr435B	(2497)	
Arg426-Gly431	(2521)	
Ile423-Met434	(2503)	
Gln422-Tyr435	(2497)	
Arg426-Lys432	(2521)	
Arg426-Gly431B	(2521)	
Asn425-Lys432	(2515)	
Consensus	(2521)	CGCGCCCTGCTGTAACTCGAG

FIG. 4M

		30
Leu122-Ser199-Tryp427-Gly431	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Val1127-Asn195-Arg426-Gly431	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Val1120-Thr202-Ile424-Ala433	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Leu122-Ser199-Arg426-Lys432	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Leu122-Ser199-Arg426-Gly431	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Lys121-Val200-Asn425-Lys432	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Val1120-Ile201-Ile424-Ala433	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Val1120-Ile201B-Ile424-Ala433	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
Consensus	(1)	GAATTCGCCACCATGGATGCAATGAAGAGA
		60
Leu122-Ser199-Tryp427-Gly431	(31)	GGGCTCTGCTGTGTGCTGCTGCTGTGTGGA
Val1127-Asn195-Arg426-Gly431	(31)	GGGCTCTGCTGTGTGCTGCTGCTGTGTGGA
Val1120-Thr202-Ile424-Ala433	(31)	GGGCTCTGCTGTGTGCTGCTGCTGTGTGGA
Leu122-Ser199-Arg426-Lys432	(31)	GGGCTCTGCTGTGTGCTGCTGCTGTGTGGA
Leu122-Ser199-Arg426-Gly431	(31)	GGGCTCTGCTGTGTGCTGCTGCTGTGTGGA
Lys121-Val200-Asn425-Lys432	(31)	GGGCTCTGCTGTGTGCTGCTGCTGTGTGGA
Val1120-Ile201-Ile424-Ala433	(31)	GGGCTCTGCTGTGTGCTGCTGCTGTGTGGA
Val1120-Ile201B-Ile424-Ala433	(31)	GGGCTCTGCTGTGTGCTGCTGCTGTGTGGA
Consensus	(31)	GGGCTCTGCTGTGTGCTGCTGCTGTGTGGA
		90
Leu122-Ser199-Tryp427-Gly431	(61)	GCAGTCTTCGTTTCGCCACGCGCCGTGGAG
Val1127-Asn195-Arg426-Gly431	(61)	GCAGTCTTCGTTTCGCCACGCGCCGTGGAG
Val1120-Thr202-Ile424-Ala433	(61)	GCAGTCTTCGTTTCGCCACGCGCCGTGGAG
Leu122-Ser199-Arg426-Lys432	(61)	GCAGTCTTCGTTTCGCCACGCGCCGTGGAG
Leu122-Ser199-Arg426-Gly431	(61)	GCAGTCTTCGTTTCGCCACGCGCCGTGGAG
Lys121-Val200-Asn425-Lys432	(61)	GCAGTCTTCGTTTCGCCACGCGCCGTGGAG
Val1120-Ile201-Ile424-Ala433	(61)	GCAGTCTTCGTTTCGCCACGCGCCGTGGAG
Val1120-Ile201B-Ile424-Ala433	(61)	GCAGTCTTCGTTTCGCCACGCGCCGTGGAG
Consensus	(61)	GCAGTCTTCGTTTCGCCACGCGCCGTGGAG
		120
Leu122-Ser199-Tryp427-Gly431	(91)	AAGCTGTGGGTGACCGTGACTACGGCGTG
Val1127-Asn195-Arg426-Gly431	(91)	AAGCTGTGGGTGACCGTGACTACGGCGTG
Val1120-Thr202-Ile424-Ala433	(91)	AAGCTGTGGGTGACCGTGACTACGGCGTG
Leu122-Ser199-Arg426-Lys432	(91)	AAGCTGTGGGTGACCGTGACTACGGCGTG
Leu122-Ser199-Arg426-Gly431	(91)	AAGCTGTGGGTGACCGTGACTACGGCGTG
Lys121-Val200-Asn425-Lys432	(91)	AAGCTGTGGGTGACCGTGACTACGGCGTG
Val1120-Ile201-Ile424-Ala433	(91)	AAGCTGTGGGTGACCGTGACTACGGCGTG
Val1120-Ile201B-Ile424-Ala433	(91)	AAGCTGTGGGTGACCGTGACTACGGCGTG
Consensus	(91)	AAGCTGTGGGTGACCGTGACTACGGCGTG
		150
Leu122-Ser199-Tryp427-Gly431	(121)	CCCGTGTGGGAAGGAGGCCACCAACCCCTG
Val1127-Asn195-Arg426-Gly431	(121)	CCCGTGTGGGAAGGAGGCCACCAACCCCTG
Val1120-Thr202-Ile424-Ala433	(121)	CCCGTGTGGGAAGGAGGCCACCAACCCCTG
Leu122-Ser199-Arg426-Lys432	(121)	CCCGTGTGGGAAGGAGGCCACCAACCCCTG
Leu122-Ser199-Arg426-Gly431	(121)	CCCGTGTGGGAAGGAGGCCACCAACCCCTG
Lys121-Val200-Asn425-Lys432	(121)	CCCGTGTGGGAAGGAGGCCACCAACCCCTG
Val1120-Ile201-Ile424-Ala433	(121)	CCCGTGTGGGAAGGAGGCCACCAACCCCTG
Val1120-Ile201B-Ile424-Ala433	(121)	CCCGTGTGGGAAGGAGGCCACCAACCCCTG
Consensus	(121)	CCCGTGTGGGAAGGAGGCCACCAACCCCTG
		180
Leu122-Ser199-Tryp427-Gly431	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Val1127-Asn195-Arg426-Gly431	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Val1120-Thr202-Ile424-Ala433	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Leu122-Ser199-Arg426-Lys432	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Leu122-Ser199-Arg426-Gly431	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC
Lys121-Val200-Asn425-Lys432	(151)	TTCTGCGCCAGCGACGCCAAGGCCTACGAC

FIG. 5A

WO 00/39303	29 / 65	PCT/US99/31272
Vall120-Ile201-Ile424-Ala433	(151) TTCTGCGCCAGCGACGCCAAGGCTACGAC	
Vall120-Ile201B-Ile424-Ala433	(151) TTCTGCGCCAGCGACGCCAAGGCTACGAC	
Consensus	(151) TTCTGCGCCAGCGACGCCAAGGCTACGAC	
Leul122-Ser199-Tryp427-Gly431	181 210	
Vall127-Asn195-Arg426-Gly431	(181) ACCGAGGTGCACAACGTGTGGGCCACCCAC	
Vall120-Thr202-Ile424-Ala433	(181) ACCGAGGTGCACAACGTGTGGGCCACCCAC	
Leul122-Ser199-Arg426-Lys432	(181) ACCGAGGTGCACAACGTGTGGGCCACCCAC	
Leul122-Ser199-Arg426-Gly431	(181) ACCGAGGTGCACAACGTGTGGGCCACCCAC	
Lys121-Val200-Asn425-Lys432	(181) ACCGAGGTGCACAACGTGTGGGCCACCCAC	
Vall120-Ile201-Ile424-Ala433	(181) ACCGAGGTGCACAACGTGTGGGCCACCCAC	
Vall120-Ile201B-Ile424-Ala433	(181) ACCGAGGTGCACAACGTGTGGGCCACCCAC	
Consensus	(181) ACCGAGGTGCACAACGTGTGGGCCACCCAC	
Leul122-Ser199-Tryp427-Gly431	211 240	
Vall127-Asn195-Arg426-Gly431	(211) GCCTGCGTGCACCGACCCCAACCCCGAG	
Vall120-Thr202-Ile424-Ala433	(211) GCCTGCGTGCACCGACCCCAACCCCGAG	
Leul122-Ser199-Arg426-Lys432	(211) GCCTGCGTGCACCGACCCCAACCCCGAG	
Leul122-Ser199-Arg426-Gly431	(211) GCCTGCGTGCACCGACCCCAACCCCGAG	
Lys121-Val200-Asn425-Lys432	(211) GCCTGCGTGCACCGACCCCAACCCCGAG	
Vall120-Ile201-Ile424-Ala433	(211) GCCTGCGTGCACCGACCCCAACCCCGAG	
Vall120-Ile201B-Ile424-Ala433	(211) GCCTGCGTGCACCGACCCCAACCCCGAG	
Consensus	(211) GCCTGCGTGCACCGACCCCAACCCCGAG	
Leul122-Ser199-Tryp427-Gly431	241 270	
Vall127-Asn195-Arg426-Gly431	(241) GAGATCGTGCTGGAGAACGTGACCGAGAAC	
Vall120-Thr202-Ile424-Ala433	(241) GAGATCGTGCTGGAGAACGTGACCGAGAAC	
Leul122-Ser199-Arg426-Lys432	(241) GAGATCGTGCTGGAGAACGTGACCGAGAAC	
Leul122-Ser199-Arg426-Gly431	(241) GAGATCGTGCTGGAGAACGTGACCGAGAAC	
Lys121-Val200-Asn425-Lys432	(241) GAGATCGTGCTGGAGAACGTGACCGAGAAC	
Vall120-Ile201-Ile424-Ala433	(241) GAGATCGTGCTGGAGAACGTGACCGAGAAC	
Vall120-Ile201B-Ile424-Ala433	(241) GAGATCGTGCTGGAGAACGTGACCGAGAAC	
Consensus	(241) GAGATCGTGCTGGAGAACGTGACCGAGAAC	
Leul122-Ser199-Tryp427-Gly431	271 300	
Vall127-Asn195-Arg426-Gly431	(271) TTCAACATGTGGAAGAACCAATGGTGGAG	
Vall120-Thr202-Ile424-Ala433	(271) TTCAACATGTGGAAGAACCAATGGTGGAG	
Leul122-Ser199-Arg426-Lys432	(271) TTCAACATGTGGAAGAACCAATGGTGGAG	
Leul122-Ser199-Arg426-Gly431	(271) TTCAACATGTGGAAGAACCAATGGTGGAG	
Lys121-Val200-Asn425-Lys432	(271) TTCAACATGTGGAAGAACCAATGGTGGAG	
Vall120-Ile201-Ile424-Ala433	(271) TTCAACATGTGGAAGAACCAATGGTGGAG	
Vall120-Ile201B-Ile424-Ala433	(271) TTCAACATGTGGAAGAACCAATGGTGGAG	
Consensus	(271) TTCAACATGTGGAAGAACCAATGGTGGAG	
Leul122-Ser199-Tryp427-Gly431	301 330	
Vall127-Asn195-Arg426-Gly431	(301) CAGATGCACGAGGACATCATCAGCCTGTGG	
Vall120-Thr202-Ile424-Ala433	(301) CAGATGCACGAGGACATCATCAGCCTGTGG	
Leul122-Ser199-Arg426-Lys432	(301) CAGATGCACGAGGACATCATCAGCCTGTGG	
Leul122-Ser199-Arg426-Gly431	(301) CAGATGCACGAGGACATCATCAGCCTGTGG	
Lys121-Val200-Asn425-Lys432	(301) CAGATGCACGAGGACATCATCAGCCTGTGG	
Vall120-Ile201-Ile424-Ala433	(301) CAGATGCACGAGGACATCATCAGCCTGTGG	
Vall120-Ile201B-Ile424-Ala433	(301) CAGATGCACGAGGACATCATCAGCCTGTGG	
Consensus	(301) CAGATGCACGAGGACATCATCAGCCTGTGG	
Leul122-Ser199-Tryp427-Gly431	331 360	
Vall127-Asn195-Arg426-Gly431	(331) GACCCAGAGCCTGAGGCCCTGCGTGAAGCTG	
Vall120-Thr202-Ile424-Ala433	(331) GACCCAGAGCCTGAGGCCCTGCGTGAAGCTG	
	(331) GACCCAGAGCCTGAGGCCCTGCGTGAAGCTG	

FIG. 5B

WO 00/39303	30 / 65	PCT/US99/31272
Leu122-Ser199-Arg426-Lys432	(331)	GACCAAGCCTGAAGCCCTGCGTGAAGCTG
Leu122-Ser199-Arg426-Gly431	(331)	GACCAGAGCCTGAAGCCCTGCGTGAAGCTG
Lys121-Val200-Asn425-Lys432	(331)	GACCAAGCCTGAAGCCCTGCGTGAAGCTG
Val120-Ile201-Ile424-Ala433	(331)	GACCAAGCCTGAAGCCCTGCGTGAAGCTG
Val120-Ile201B-Ile424-Ala433	(331)	GACCAAGCCTGAAGCCCTGCGTGAAGCTG
Consensus	(331)	GACCAAGCCTGAAGCCCTGCGTGAAGCTG
Leu122-Ser199-Trp427-Gly431	(361)	-----GG-----
Val127-Asn195-Arg426-Gly431	(361)	ACCCCCCTGTGCGTGGGGGAGGGAAGCTG
Val120-Thr202-Ile424-Ala433	(355)	-----GG-----
Leu122-Ser199-Arg426-Lys432	(361)	-----GG-----
Leu122-Ser199-Arg426-Gly431	(361)	-----GG-----
Lys121-Val200-Asn425-Lys432	(357)	-----GG-----
Val120-Ile201-Ile424-Ala433	(355)	-----
Val120-Ile201B-Ile424-Ala433	(355)	-----
Consensus	(361)	GG
Leu122-Ser199-Trp427-Gly431	(363)	391 ----- 420
Val127-Asn195-Arg426-Gly431	(391)	---CAACAGCGTGATCACCCAGGCGCTGCCCC
Val120-Thr202-Ile424-Ala433	(357)	AACACCAAGCGTGATCACCCAGGCGCTGCCCC
Leu122-Ser199-Arg426-Lys432	(363)	-----CGGCGC---CACCCAGGCGCTGCCCC
Leu122-Ser199-Arg426-Gly431	(363)	---CAACAGCGTGATCACCCAGGCGCTGCCCC
Lys121-Val200-Asn425-Lys432	(359)	---CAACAGCGTGATCACCCAGGCGCTGCCCC
Val120-Ile201-Ile424-Ala433	(355)	-----CCCCGTGATCACCCAGGCGCTGCCCC
Val120-Ile201B-Ile424-Ala433	(355)	-----CGGCGCATCACCCAGGCGCTGCCCC
Consensus	(391)	CA CAGCGTGATCACCCAGGCGCTGCCCC
Leu122-Ser199-Trp427-Gly431	(391)	421 ----- 450
Val127-Asn195-Arg426-Gly431	(421)	AAGGTGAGCTTCGAGCCCATCCCATCCAC
Val120-Thr202-Ile424-Ala433	(379)	AAGGTGAGCTTCGAGCCCATCCCATCCAC
Leu122-Ser199-Arg426-Lys432	(391)	AAGGTGAGCTTCGAGCCCATCCCATCCAC
Leu122-Ser199-Arg426-Gly431	(391)	AAGGTGAGCTTCGAGCCCATCCCATCCAC
Lys121-Val200-Asn425-Lys432	(385)	AAGGTGAGCTTCGAGCCCATCCCATCCAC
Val120-Ile201-Ile424-Ala433	(379)	AAGGTGAGCTTCGAGCCCATCCCATCCAC
Val120-Ile201B-Ile424-Ala433	(379)	AAGGTGAGCTTCGAGCCCATCCCATCCAC
Consensus	(421)	AAGGTGAGCTTCGAGCCCATCCCATCCAC
Leu122-Ser199-Trp427-Gly431	(421)	451 ----- 480
Val127-Asn195-Arg426-Gly431	(451)	TACTGCGCCCGCGCGGCTTCGCCATCCTG
Val120-Thr202-Ile424-Ala433	(409)	TACTGCGCCCGCGCGGCTTCGCCATCCTG
Leu122-Ser199-Arg426-Lys432	(421)	TACTGCGCCCGCGCGGCTTCGCCATCCTG
Leu122-Ser199-Arg426-Gly431	(421)	TACTGCGCCCGCGCGGCTTCGCCATCCTG
Lys121-Val200-Asn425-Lys432	(415)	TACTGCGCCCGCGCGGCTTCGCCATCCTG
Val120-Ile201-Ile424-Ala433	(409)	TACTGCGCCCGCGCGGCTTCGCCATCCTG
Val120-Ile201B-Ile424-Ala433	(409)	TACTGCGCCCGCGCGGCTTCGCCATCCTG
Consensus	(451)	TACTGCGCCCGCGCGGCTTCGCCATCCTG
Leu122-Ser199-Trp427-Gly431	(451)	481 ----- 510
Val127-Asn195-Arg426-Gly431	(481)	AAGTCCAAACGACAGAAGTTCAACGGGAGC
Val120-Thr202-Ile424-Ala433	(439)	AAGTCCAAACGACAGAAGTTCAACGGGAGC
Leu122-Ser199-Arg426-Lys432	(451)	AAGTCCAAACGACAGAAGTTCAACGGGAGC
Leu122-Ser199-Arg426-Gly431	(451)	AAGTCCAAACGACAGAAGTTCAACGGGAGC
Lys121-Val200-Asn425-Lys432	(445)	AAGTCCAAACGACAGAAGTTCAACGGGAGC
Val120-Ile201-Ile424-Ala433	(439)	AAGTCCAAACGACAGAAGTTCAACGGGAGC
Val120-Ile201B-Ile424-Ala433	(439)	AAGTCCAAACGACAGAAGTTCAACGGGAGC
Consensus	(481)	AAGTCCAAACGACAGAAGTTCAACGGGAGC
		511 ----- 540

FIG. 5C

WO 00/39303	31 / 65	PCT/US99/31272
Leu122-Ser199-Tryp427-Gly431	(481)	GGCCCCGACCAACGCTGAGCACCGCTGGAG
Val1127-Asn195-Arg426-Gly431	(511)	GGCCCCGACCAACGCTGAGCACCGCTGGAG
Val1120-Thr202-Ile424-Ala433	(469)	GGCCCCGACCAACGCTGAGCACCGCTGGAG
Leu122-Ser199-Arg426-Lys432	(481)	GGCCCCGACCAACGCTGAGCACCGCTGGAG
Leu122-Ser199-Arg426-Gly431	(481)	GGCCCCGACCAACGCTGAGCACCGCTGGAG
Lys121-Val200-Asn425-Lys432	(475)	GGCCCCGACCAACGCTGAGCACCGCTGGAG
Val1120-Ile201-Ile424-Ala433	(469)	GGCCCCGACCAACGCTGAGCACCGCTGGAG
Val1120-Ile201B-Ile424-Ala433	(469)	GGCCCCGACCAACGCTGAGCACCGCTGGAG
Consensus	(511)	GGCCCCGACCAACGCTGAGCACCGCTGGAG
Leu122-Ser199-Tryp427-Gly431	541	570
Val1127-Asn195-Arg426-Gly431	(511)	TGCACCCACGGCATCCGCCCGCTGGTGAGC
Val1120-Thr202-Ile424-Ala433	(499)	TGCACCCACGGCATCCGCCCGCTGGTGAGC
Leu122-Ser199-Arg426-Lys432	(511)	TGCACCCACGGCATCCGCCCGCTGGTGAGC
Leu122-Ser199-Arg426-Gly431	(511)	TGCACCCACGGCATCCGCCCGCTGGTGAGC
Lys121-Val200-Asn425-Lys432	(505)	TGCACCCACGGCATCCGCCCGCTGGTGAGC
Val1120-Ile201-Ile424-Ala433	(499)	TGCACCCACGGCATCCGCCCGCTGGTGAGC
Val1120-Ile201B-Ile424-Ala433	(499)	TGCACCCACGGCATCCGCCCGCTGGTGAGC
Consensus	(541)	TGCACCCACGGCATCCGCCCGCTGGTGAGC
Leu122-Ser199-Tryp427-Gly431	(541)	571 600
Val1127-Asn195-Arg426-Gly431	(571)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Val1120-Thr202-Ile424-Ala433	(529)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Leu122-Ser199-Arg426-Lys432	(541)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Leu122-Ser199-Arg426-Gly431	(541)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Lys121-Val200-Asn425-Lys432	(535)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Val1120-Ile201-Ile424-Ala433	(529)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Val1120-Ile201B-Ile424-Ala433	(529)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Consensus	(571)	ACCCAGCTGCTGCTGAACGGCAGCCTGGCC
Leu122-Ser199-Tryp427-Gly431	(571)	601 630
Val1127-Asn195-Arg426-Gly431	(601)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Val1120-Thr202-Ile424-Ala433	(559)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Leu122-Ser199-Arg426-Lys432	(571)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Leu122-Ser199-Arg426-Gly431	(571)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Lys121-Val200-Asn425-Lys432	(565)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Val1120-Ile201-Ile424-Ala433	(559)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Val1120-Ile201B-Ile424-Ala433	(559)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Consensus	(601)	GAGGAGGGCGTGGTGATCCGCAGCGAGAAC
Leu122-Ser199-Tryp427-Gly431	(601)	631 660
Val1127-Asn195-Arg426-Gly431	(631)	TTCCACCGACAAAGCCAAAGACCATCATCGTG
Val1120-Thr202-Ile424-Ala433	(589)	TTCCACCGACAAAGCCAAAGACCATCATCGTG
Leu122-Ser199-Arg426-Lys432	(601)	TTCCACCGACAAAGCCAAAGACCATCATCGTG
Leu122-Ser199-Arg426-Gly431	(601)	TTCCACCGACAAAGCCAAAGACCATCATCGTG
Lys121-Val200-Asn425-Lys432	(595)	TTCCACCGACAAAGCCAAAGACCATCATCGTG
Val1120-Ile201-Ile424-Ala433	(589)	TTCCACCGACAAAGCCAAAGACCATCATCGTG
Val1120-Ile201B-Ile424-Ala433	(589)	TTCCACCGACAAAGCCAAAGACCATCATCGTG
Consensus	(631)	TTCCACCGACAAAGCCAAAGACCATCATCGTG
Leu122-Ser199-Tryp427-Gly431	(631)	661 690
Val1127-Asn195-Arg426-Gly431	(661)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Val1120-Thr202-Ile424-Ala433	(619)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Leu122-Ser199-Arg426-Lys432	(631)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Leu122-Ser199-Arg426-Gly431	(631)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Lys121-Val200-Asn425-Lys432	(625)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC
Val1120-Ile201-Ile424-Ala433	(619)	CAGCTGAAGGAGAGCGTGGAGATCAACTGC

FIG. 5D

WO 00/39303	32 / 65	PCT/US99/31272
Val120-Ile201B-Ile424-Ala433	(619) CAGCTGAAGGAGAGCGTGGAGATCAACTGC	
Consensus	(661) CAGCTGAAGGAGAGCGTGGAGATCAACTGC	691 720
Leu122-Ser199-Tryp427-Gly431	(661) ACCCGCCCCAACAAACACCCGCAAGAGC	
Val127-Asn195-Arg426-Gly431	(691) ACCCGCCCCAACAAACACCCGCAAGAGC	
Val120-Thr202-Ile424-Ala433	(649) ACCCGCCCCAACAAACACCCGCAAGAGC	
Leu122-Ser199-Arg426-Lys432	(661) ACCCGCCCCAACAAACACCCGCAAGAGC	
Leu122-Ser199-Arg426-Gly431	(661) ACCCGCCCCAACAAACACCCGCAAGAGC	
Lys121-Val200-Asn425-Lys432	(655) ACCCGCCCCAACAAACACCCGCAAGAGC	
Val120-Ile201-Ile424-Ala433	(649) ACCCGCCCCAACAAACACCCGCAAGAGC	
Val120-Ile201B-Ile424-Ala433	(649) ACCCGCCCCAACAAACACCCGCAAGAGC	
Consensus	(691) ACCCGCCCCAACAAACACCCGCAAGAGC	721 750
Leu122-Ser199-Tryp427-Gly431	(691) ATCACCATCGGCCCGGCCGCGCTTCTAC	
Val127-Asn195-Arg426-Gly431	(721) ATCACCATCGGCCCGGCCGCGCTTCTAC	
Val120-Thr202-Ile424-Ala433	(679) ATCACCATCGGCCCGGCCGCGCTTCTAC	
Leu122-Ser199-Arg426-Lys432	(691) ATCACCATCGGCCCGGCCGCGCTTCTAC	
Leu122-Ser199-Arg426-Gly431	(691) ATCACCATCGGCCCGGCCGCGCTTCTAC	
Lys121-Val200-Asn425-Lys432	(685) ATCACCATCGGCCCGGCCGCGCTTCTAC	
Val120-Ile201-Ile424-Ala433	(679) ATCACCATCGGCCCGGCCGCGCTTCTAC	
Val120-Ile201B-Ile424-Ala433	(679) ATCACCATCGGCCCGGCCGCGCTTCTAC	
Consensus	(721) ATCACCATCGGCCCGGCCGCGCTTCTAC	751 780
Leu122-Ser199-Tryp427-Gly431	(721) GCCACCGGGGACATCATCGGCACATCCGC	
Val127-Asn195-Arg426-Gly431	(751) GCCACCGGGGACATCATCGGCACATCCGC	
Val120-Thr202-Ile424-Ala433	(709) GCCACCGGGGACATCATCGGCACATCCGC	
Leu122-Ser199-Arg426-Lys432	(721) GCCACCGGGGACATCATCGGCACATCCGC	
Leu122-Ser199-Arg426-Gly431	(721) GCCACCGGGGACATCATCGGCACATCCGC	
Lys121-Val200-Asn425-Lys432	(715) GCCACCGGGGACATCATCGGCACATCCGC	
Val120-Ile201-Ile424-Ala433	(709) GCCACCGGGGACATCATCGGCACATCCGC	
Val120-Ile201B-Ile424-Ala433	(709) GCCACCGGGGACATCATCGGCACATCCGC	
Consensus	(751) GCCACCGGGGACATCATCGGCACATCCGC	781 810
Leu122-Ser199-Tryp427-Gly431	(751) CAGGCCCACTGCACATCAGCGGCGAAGG	
Val127-Asn195-Arg426-Gly431	(781) CAGGCCCACTGCACATCAGCGGCGAAGG	
Val120-Thr202-Ile424-Ala433	(739) CAGGCCCACTGCACATCAGCGGCGAAGG	
Leu122-Ser199-Arg426-Lys432	(751) CAGGCCCACTGCACATCAGCGGCGAAGG	
Leu122-Ser199-Arg426-Gly431	(751) CAGGCCCACTGCACATCAGCGGCGAAGG	
Lys121-Val200-Asn425-Lys432	(745) CAGGCCCACTGCACATCAGCGGCGAAGG	
Val120-Ile201-Ile424-Ala433	(739) CAGGCCCACTGCACATCAGCGGCGAAGG	
Val120-Ile201B-Ile424-Ala433	(739) CAGGCCCACTGCACATCAGCGGCGAAGG	
Consensus	(781) CAGGCCCACTGCACATCAGCGGCGAAGG	811 840
Leu122-Ser199-Tryp427-Gly431	(781) TGGACACACCCCTGAAGCAGATCGTGACC	
Val127-Asn195-Arg426-Gly431	(811) TGGACACACCCCTGAAGCAGATCGTGACC	
Val120-Thr202-Ile424-Ala433	(769) TGGACACACCCCTGAAGCAGATCGTGACC	
Leu122-Ser199-Arg426-Lys432	(781) TGGACACACCCCTGAAGCAGATCGTGACC	
Leu122-Ser199-Arg426-Gly431	(781) TGGACACACCCCTGAAGCAGATCGTGACC	
Lys121-Val200-Asn425-Lys432	(775) TGGACACACCCCTGAAGCAGATCGTGACC	
Val120-Ile201-Ile424-Ala433	(769) TGGACACACCCCTGAAGCAGATCGTGACC	
Val120-Ile201B-Ile424-Ala433	(769) TGGACACACCCCTGAAGCAGATCGTGACC	
Consensus	(811) TGGACACACCCCTGAAGCAGATCGTGACC	841 870
Leu122-Ser199-Tryp427-Gly431	(811) AAGCTGCAGGCCAGTTTCGGCAACAGACC	
Val127-Asn195-Arg426-Gly431	(841) AAGCTGCAGGCCAGTTTCGGCAACAGACC	
Val120-Thr202-Ile424-Ala433	(799) AAGCTGCAGGCCAGTTTCGGCAACAGACC	
Leu122-Ser199-Arg426-Lys432	(811) AAGCTGCAGGCCAGTTTCGGCAACAGACC	

FIG. 5E

WO 00/39303	33 / 65	PCT/US99/31272
Leu122-Ser199-Arg426-Gly431	(811)	AAGCTGCAGGCCAGTTCGGCAACAGACC
Lys121-Val200-Asn425-Lys432	(805)	AAGCTGCAGGCCAGTTCGGCAACAGACC
Val120-Ile201-Ile424-Ala433	(799)	AAGCTGCAGGCCAGTTCGGCAACAGACC
Val120-Ile201B-Ile424-Ala433	(799)	AAGCTGCAGGCCAGTTCGGCAACAGACC
Consensus	(841)	AAGCTGCAGGCCAGTTCGGCAACAGACC
Leu122-Ser199-Tryp427-Gly431	871	900
Val127-Asn195-Arg426-Gly431	(841)	ATCGTGTTCAGGCAGAGCAGCGCGGGCGAC
Val120-Thr202-Ile424-Ala433	(871)	ATCGTGTTCAGGCAGAGCAGCGCGGGCGAC
Leu122-Ser199-Arg426-Lys432	(829)	ATCGTGTTCAGGCAGAGCAGCGCGGGCGAC
Leu122-Ser199-Arg426-Gly431	(841)	ATCGTGTTCAGGCAGAGCAGCGCGGGCGAC
Lys121-Val200-Asn425-Lys432	(841)	ATCGTGTTCAGGCAGAGCAGCGCGGGCGAC
Val120-Ile201-Ile424-Ala433	(835)	ATCGTGTTCAGGCAGAGCAGCGCGGGCGAC
Val120-Ile201B-Ile424-Ala433	(829)	ATCGTGTTCAGGCAGAGCAGCGCGGGCGAC
Consensus	(829)	ATCGTGTTCAGGCAGAGCAGCGCGGGCGAC
Leu122-Ser199-Tryp427-Gly431	(871)	ATCGTGTTCAGGCAGAGCAGCGCGGGCGAC
Val127-Asn195-Arg426-Gly431	(871)	ATCGTGTTCAGGCAGAGCAGCGCGGGCGAC
Val120-Thr202-Ile424-Ala433	(901)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Leu122-Ser199-Arg426-Lys432	(859)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Leu122-Ser199-Arg426-Gly431	(871)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Lys121-Val200-Asn425-Lys432	(871)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Val120-Ile201-Ile424-Ala433	(865)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Val120-Ile201B-Ile424-Ala433	(859)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Consensus	(859)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Leu122-Ser199-Tryp427-Gly431	(901)	CCCGAGATCGTGATGCACAGCTTCAACTGC
Val127-Asn195-Arg426-Gly431	931	960
Val120-Thr202-Ile424-Ala433	(901)	GGCGGCGAGTTCCTTCTACTGCAACAGCACC
Leu122-Ser199-Arg426-Lys432	(931)	GGCGGCGAGTTCCTTCTACTGCAACAGCACC
Leu122-Ser199-Arg426-Gly431	(889)	GGCGGCGAGTTCCTTCTACTGCAACAGCACC
Lys121-Val200-Asn425-Lys432	(901)	GGCGGCGAGTTCCTTCTACTGCAACAGCACC
Val120-Ile201-Ile424-Ala433	(901)	GGCGGCGAGTTCCTTCTACTGCAACAGCACC
Val120-Ile201B-Ile424-Ala433	(895)	GGCGGCGAGTTCCTTCTACTGCAACAGCACC
Consensus	(889)	GGCGGCGAGTTCCTTCTACTGCAACAGCACC
Leu122-Ser199-Tryp427-Gly431	(889)	GGCGGCGAGTTCCTTCTACTGCAACAGCACC
Val127-Asn195-Arg426-Gly431	(931)	GGCGGCGAGTTCCTTCTACTGCAACAGCACC
Val120-Thr202-Ile424-Ala433	961	990
Leu122-Ser199-Arg426-Lys432	(931)	CAGCTGTTCACAGCAGCTTGGAAACAGACC
Leu122-Ser199-Arg426-Gly431	(961)	CAGCTGTTCACAGCAGCTTGGAAACAGACC
Val120-Thr202-Ile424-Ala433	(919)	CAGCTGTTCACAGCAGCTTGGAAACAGACC
Leu122-Ser199-Arg426-Lys432	(931)	CAGCTGTTCACAGCAGCTTGGAAACAGACC
Leu122-Ser199-Arg426-Gly431	(931)	CAGCTGTTCACAGCAGCTTGGAAACAGACC
Lys121-Val200-Asn425-Lys432	(925)	CAGCTGTTCACAGCAGCTTGGAAACAGACC
Val120-Ile201-Ile424-Ala433	(919)	CAGCTGTTCACAGCAGCTTGGAAACAGACC
Val120-Ile201B-Ile424-Ala433	(919)	CAGCTGTTCACAGCAGCTTGGAAACAGACC
Consensus	(961)	CAGCTGTTCACAGCAGCTTGGAAACAGACC
Leu122-Ser199-Tryp427-Gly431	991	1020
Val127-Asn195-Arg426-Gly431	(961)	ATCGGCCCCACACACACCAACGGCACCATC
Val120-Thr202-Ile424-Ala433	(991)	ATCGGCCCCACACACACCAACGGCACCATC
Leu122-Ser199-Arg426-Lys432	(949)	ATCGGCCCCACACACACCAACGGCACCATC
Leu122-Ser199-Arg426-Gly431	(961)	ATCGGCCCCACACACACCAACGGCACCATC
Lys121-Val200-Asn425-Lys432	(961)	ATCGGCCCCACACACACCAACGGCACCATC
Val120-Ile201-Ile424-Ala433	(955)	ATCGGCCCCACACACACCAACGGCACCATC
Val120-Ile201B-Ile424-Ala433	(949)	ATCGGCCCCACACACACCAACGGCACCATC
Consensus	(949)	ATCGGCCCCACACACACCAACGGCACCATC
Leu122-Ser199-Tryp427-Gly431	(991)	ATCGGCCCCACACACACCAACGGCACCATC
	1021	1050
	(991)	ACCGTGCCTGCGCGATCAAGCAGATCATC

FIG. 5F



Val1127-Asn195-Arg426-Gly431	(1021)	ACCCCTGCCCTGCCGCATCAAGCAGATCATC
Val1120-Thr202-Ile424-Ala433	(979)	ACCCCTGCCCTGCCGCATCAAGCAGATCATC
Leu122-Ser199-Arg426-Lys432	(991)	ACCCCTGCCCTGCCGCATCAAGCAGATCATC
Leu122-Ser199-Arg426-Gly431	(991)	ACCCCTGCCCTGCCGCATCAAGCAGATCATC
Lys121-Val200-Asn425-Lys432	(985)	ACCCCTGCCCTGCCGCATCAAGCAGATCATC
Val1120-Ile201-Ile424-Ala433	(979)	ACCCCTGCCCTGCCGCATCAAGCAGATCATC
Val1120-Ile201B-Ile424-Ala433	(979)	ACCCCTGCCCTGCCGCATCAAGCAGATCATC
Consensus	(1021)	ACCCCTGCCCTGCCGCATCAAGCAGATCATC
Leu122-Ser199 Tryp427-Gly431	(1021)	AACCCGCTGGGGGGCAAGGCCATGTACGCC
Val1127-Asn195-Arg426-Gly431	(1051)	AACCCGCGGGGGGGCAAGGCCATGTACGCC
Val1120-Thr202-Ile424-Ala433	(1009)	-----GGGGG---GCCATGTACGCC
Leu122-Ser199-Arg426-Lys432	(1021)	AACCCGCGGGGGCAACAGGCCATGTACGCC
Leu122-Ser199-Arg426-Gly431	(1021)	AACCCGCGGGGGGGCAAGGCCATGTACGCC
Lys121-Val200-Asn425-Lys432	(1015)	AAC-----GCCCGCAAGGCCATGTACGCC
Val1120-Ile201-Ile424-Ala433	(1009)	-----GGGGG---GCCATGTACGCC
Val1120-Ile201B-Ile424-Ala433	(1009)	-----GGGGG---GCCATGTACGCC
Consensus	(1051)	AACCGC G GGGGCAAGGCCATGTACGCC
Leu122-Ser199 Tryp427-Gly431	(1051)	CCGCCCATCCGCGGCCAGATCCGCTGCAGC
Val1127-Asn195-Arg426-Gly431	(1081)	CCGCCCATCCGCGGCCAGATCCGCTGCAGC
Val1120-Thr202-Ile424-Ala433	(1027)	CCGCCCATCCGCGGCCAGATCCGCTGCAGC
Leu122-Ser199-Arg426-Lys432	(1051)	CCGCCCATCCGCGGCCAGATCCGCTGCAGC
Leu122-Ser199-Arg426-Gly431	(1051)	CCGCCCATCCGCGGCCAGATCCGCTGCAGC
Lys121-Val200-Asn425-Lys432	(1039)	CCGCCCATCCGCGGCCAGATCCGCTGCAGC
Val1120-Ile201-Ile424-Ala433	(1027)	CCGCCCATCCGCGGCCAGATCCGCTGCAGC
Val1120-Ile201B-Ile424-Ala433	(1027)	CCGCCCATCCGCGGCCAGATCCGCTGCAGC
Consensus	(1081)	CCGCCCATCCGCGGCCAGATCCGCTGCAGC
Leu122-Ser199 Tryp427-Gly431	(1081)	AGCAACATCACCGGCCTGCTGTCAGCCGC
Val1127-Asn195-Arg426-Gly431	(1111)	AGCAACATCACCGGCCTGCTGTCAGCCGC
Val1120-Thr202-Ile424-Ala433	(1057)	AGCAACATCACCGGCCTGCTGTCAGCCGC
Leu122-Ser199-Arg426-Lys432	(1081)	AGCAACATCACCGGCCTGCTGTCAGCCGC
Leu122-Ser199-Arg426-Gly431	(1081)	AGCAACATCACCGGCCTGCTGTCAGCCGC
Lys121-Val200-Asn425-Lys432	(1069)	AGCAACATCACCGGCCTGCTGTCAGCCGC
Val1120-Ile201-Ile424-Ala433	(1057)	AGCAACATCACCGGCCTGCTGTCAGCCGC
Val1120-Ile201B-Ile424-Ala433	(1057)	AGCAACATCACCGGCCTGCTGTCAGCCGC
Consensus	(1111)	AGCAACATCACCGGCCTGCTGTCAGCCGC
Leu122-Ser199 Tryp427-Gly431	(1111)	GACGCGGGCAAGGAGATCAGCAACACACC
Val1127-Asn195-Arg426-Gly431	(1141)	GACGCGGGCAAGGAGATCAGCAACACACC
Val1120-Thr202-Ile424-Ala433	(1087)	GACGCGGGCAAGGAGATCAGCAACACACC
Leu122-Ser199-Arg426-Lys432	(1111)	GACGCGGGCAAGGAGATCAGCAACACACC
Leu122-Ser199-Arg426-Gly431	(1111)	GACGCGGGCAAGGAGATCAGCAACACACC
Lys121-Val200-Asn425-Lys432	(1099)	GACGCGGGCAAGGAGATCAGCAACACACC
Val1120-Ile201-Ile424-Ala433	(1087)	GACGCGGGCAAGGAGATCAGCAACACACC
Val1120-Ile201B-Ile424-Ala433	(1087)	GACGCGGGCAAGGAGATCAGCAACACACC
Consensus	(1141)	GACGCGGGCAAGGAGATCAGCAACACACC
Leu122-Ser199 Tryp427-Gly431	(1141)	GAGATCTTCCGCCCGGGCGGGCGGACATG
Val1127-Asn195-Arg426-Gly431	(1171)	GAGATCTTCCGCCCGGGCGGGCGGACATG
Val1120-Thr202-Ile424-Ala433	(1117)	GAGATCTTCCGCCCGGGCGGGCGGACATG
Leu122-Ser199-Arg426-Lys432	(1141)	GAGATCTTCCGCCCGGGCGGGCGGACATG
Leu122-Ser199-Arg426-Gly431	(1141)	GAGATCTTCCGCCCGGGCGGGCGGACATG
Lys121-Val200-Asn425-Lys432	(1129)	GAGATCTTCCGCCCGGGCGGGCGGACATG
Val1120-Ile201-Ile424-Ala433	(1117)	GAGATCTTCCGCCCGGGCGGGCGGACATG
Val1120-Ile201B-Ile424-Ala433	(1117)	GAGATCTTCCGCCCGGGCGGGCGGACATG

FIG. 5G

Consensus	(1171)	GAGATCTTCCGCCCGCGCGGGCGACATG
	1201	1230
Leu122-Ser199 Tryp427-Gly431	(1171)	CGCGCAACCTGGGSCAGCGAGCTGTACAAAG
Val1127-Asn195-Arg426-Gly431	(1201)	CGCGCAACCTGGGSCAGCGAGCTGTACAAAG
Val1120-Thr202-Ile424-Ala433	(1147)	CGCGCAACCTGGGSCAGCGAGCTGTACAAAG
Leu122-Ser199-Arg426-Lys432	(1171)	CGCGCAACCTGGGSCAGCGAGCTGTACAAAG
Leu122-Ser199-Arg426-Gly431	(1171)	CGCGCAACCTGGGSCAGCGAGCTGTACAAAG
Lys121-Val200-Asn425-Lys432	(1159)	CGCGCAACCTGGGSCAGCGAGCTGTACAAAG
Val1120-Ile201-Ile424-Ala433	(1147)	CGCGCAACCTGGGSCAGCGAGCTGTACAAAG
Val1120-Ile201B-Ile424-Ala433	(1147)	CGCGCAACCTGGGSCAGCGAGCTGTACAAAG
Consensus	(1201)	CGCGCAACCTGGGSCAGCGAGCTGTACAAAG
	1231	1260
Leu122-Ser199 Tryp427-Gly431	(1201)	TACAAAGTGTGGAAGATCGAGCCCTGGGC
Val1127-Asn195-Arg426-Gly431	(1231)	TACAAAGTGTGGAAGATCGAGCCCTGGGC
Val1120-Thr202-Ile424-Ala433	(1177)	TACAAAGTGTGGAAGATCGAGCCCTGGGC
Leu122-Ser199-Arg426-Lys432	(1201)	TACAAAGTGTGGAAGATCGAGCCCTGGGC
Leu122-Ser199-Arg426-Gly431	(1201)	TACAAAGTGTGGAAGATCGAGCCCTGGGC
Lys121-Val200-Asn425-Lys432	(1189)	TACAAAGTGTGGAAGATCGAGCCCTGGGC
Val1120-Ile201-Ile424-Ala433	(1177)	TACAAAGTGTGGAAGATCGAGCCCTGGGC
Val1120-Ile201B-Ile424-Ala433	(1177)	TACAAAGTGTGGAAGATCGAGCCCTGGGC
Consensus	(1231)	TACAAAGTGTGGAAGATCGAGCCCTGGGC
	1261	1290
Leu122-Ser199 Tryp427-Gly431	(1231)	GTGGCCCCCAAGAGGCCAAGCGCGCGTG
Val1127-Asn195-Arg426-Gly431	(1261)	GTGGCCCCCAAGAGGCCAAGCGCGCGTG
Val1120-Thr202-Ile424-Ala433	(1207)	GTGGCCCCCAAGAGGCCAAGCGCGCGTG
Leu122-Ser199-Arg426-Lys432	(1231)	GTGGCCCCCAAGAGGCCAAGCGCGCGTG
Leu122-Ser199-Arg426-Gly431	(1231)	GTGGCCCCCAAGAGGCCAAGCGCGCGTG
Lys121-Val200-Asn425-Lys432	(1219)	GTGGCCCCCAAGAGGCCAAGCGCGCGTG
Val1120-Ile201-Ile424-Ala433	(1207)	GTGGCCCCCAAGAGGCCAAGCGCGCGTG
Val1120-Ile201B-Ile424-Ala433	(1207)	GTGGCCCCCAAGAGGCCAAGCGCGCGTG
Consensus	(1261)	GTGGCCCCCAAGAGGCCAAGCGCGCGTG
	1291	1320
Leu122-Ser199 Tryp427-Gly431	(1261)	GTGCAGCGCGAGAGCGCGCGGTGACCGTG
Val1127-Asn195-Arg426-Gly431	(1291)	GTGCAGCGCGAGAGCGCGCGGTGACCGTG
Val1120-Thr202-Ile424-Ala433	(1237)	GTGCAGCGCGAGAGCGCGCGGTGACCGTG
Leu122-Ser199-Arg426-Lys432	(1261)	GTGCAGCGCGAGAGCGCGCGGTGACCGTG
Leu122-Ser199-Arg426-Gly431	(1261)	GTGCAGCGCGAGAGCGCGCGGTGACCGTG
Lys121-Val200-Asn425-Lys432	(1249)	GTGCAGCGCGAGAGCGCGCGGTGACCGTG
Val1120-Ile201-Ile424-Ala433	(1237)	GTGCAGCGCGAGAGCGCGCGGTGACCGTG
Val1120-Ile201B-Ile424-Ala433	(1237)	GTGCAGCGCGAGAGCGCGCGGTGACCGTG
Consensus	(1291)	GTGCAGCGCGAGAGCGCGCGGTGACCGTG
	1321	1350
Leu122-Ser199 Tryp427-Gly431	(1291)	GGCGGCATGTCCTGGGCTTCCTGGGCGCC
Val1127-Asn195-Arg426-Gly431	(1321)	GGCGGCATGTCCTGGGCTTCCTGGGCGCC
Val1120-Thr202-Ile424-Ala433	(1267)	GGCGGCATGTCCTGGGCTTCCTGGGCGCC
Leu122-Ser199-Arg426-Lys432	(1291)	GGCGGCATGTCCTGGGCTTCCTGGGCGCC
Leu122-Ser199-Arg426-Gly431	(1291)	GGCGGCATGTCCTGGGCTTCCTGGGCGCC
Lys121-Val200-Asn425-Lys432	(1279)	GGCGGCATGTCCTGGGCTTCCTGGGCGCC
Val1120-Ile201-Ile424-Ala433	(1267)	GGCGGCATGTCCTGGGCTTCCTGGGCGCC
Val1120-Ile201B-Ile424-Ala433	(1267)	GGCGGCATGTCCTGGGCTTCCTGGGCGCC
Consensus	(1321)	GGCGGCATGTCCTGGGCTTCCTGGGCGCC
	1351	1380
Leu122-Ser199 Tryp427-Gly431	(1321)	GGCGGCATGTCCTGGGCTTCCTGGGCGCC
Val1127-Asn195-Arg426-Gly431	(1351)	GGCGGCATGTCCTGGGCTTCCTGGGCGCC
Val1120-Thr202-Ile424-Ala433	(1297)	GGCGGCATGTCCTGGGCTTCCTGGGCGCC
Leu122-Ser199-Arg426-Lys432	(1321)	GGCGGCATGTCCTGGGCTTCCTGGGCGCC
Leu122-Ser199-Arg426-Gly431	(1321)	GGCGGCATGTCCTGGGCTTCCTGGGCGCC

FIG. 5H

Lys121-Val200-Asn425-Lys432  
 Val120-Ile201-Ile424-Ala433  
 Val120-Ile201B-Ile424-Ala433  
 Consensus

Leu122-Ser199-Tryp427-Gly431  
 Val127-Asn195-Arg426-Gly431  
 Val120-Thr202-Ile424-Ala433  
 Leu122-Ser199-Arg426-Lys432  
 Leu122-Ser199-Arg426-Gly431  
 Lys121-Val200-Asn425-Lys432  
 Val120-Ile201-Ile424-Ala433  
 Val120-Ile201B-Ile424-Ala433  
 Consensus

Leu122-Ser199-Tryp427-Gly431  
 Val127-Asn195-Arg426-Gly431  
 Val120-Thr202-Ile424-Ala433  
 Leu122-Ser199-Arg426-Lys432  
 Leu122-Ser199-Arg426-Gly431  
 Lys121-Val200-Asn425-Lys432  
 Val120-Ile201-Ile424-Ala433  
 Val120-Ile201B-Ile424-Ala433  
 Consensus

Leu122-Ser199-Tryp427-Gly431  
 Val127-Asn195-Arg426-Gly431  
 Val120-Thr202-Ile424-Ala433  
 Leu122-Ser199-Arg426-Lys432  
 Leu122-Ser199-Arg426-Gly431  
 Lys121-Val200-Asn425-Lys432  
 Val120-Ile201-Ile424-Ala433  
 Val120-Ile201B-Ile424-Ala433  
 Consensus

Leu122-Ser199-Tryp427-Gly431  
 Val127-Asn195-Arg426-Gly431  
 Val120-Thr202-Ile424-Ala433  
 Leu122-Ser199-Arg426-Lys432  
 Leu122-Ser199-Arg426-Gly431  
 Lys121-Val200-Asn425-Lys432  
 Val120-Ile201-Ile424-Ala433  
 Val120-Ile201B-Ile424-Ala433  
 Consensus

Leu122-Ser199-Tryp427-Gly431  
 Val127-Asn195-Arg426-Gly431  
 Val120-Thr202-Ile424-Ala433  
 Leu122-Ser199-Arg426-Lys432  
 Leu122-Ser199-Arg426-Gly431  
 Lys121-Val200-Asn425-Lys432  
 Val120-Ile201-Ile424-Ala433  
 Val120-Ile201B-Ile424-Ala433  
 Consensus

Leu122-Ser199-Tryp427-Gly431  
 Val127-Asn195-Arg426-Gly431

(1309) GCCGCGAGCACCATGGGCGCCCGCAGCCTG  
 (1297) GCCGCGAGCACCATGGGCGCCCGCAGCCTG  
 (1297) GCCGCGAGCACCATGGGCGCCCGCAGCCTG  
 (1351) GCCGCGAGCACCATGGGCGCCCGCAGCCTG

1381 1410  
 (1351) ACCCTGACCGTGCAGGCCCGCCAGCTGCTG  
 (1381) ACCCTGACCGTGCAGGCCCGCCAGCTGCTG  
 (1327) ACCCTGACCGTGCAGGCCCGCCAGCTGCTG  
 (1352) ACCCTGACCGTGCAGGCCCGCCAGCTGCTG  
 (1351) ACCCTGACCGTGCAGGCCCGCCAGCTGCTG  
 (1339) ACCCTGACCGTGCAGGCCCGCCAGCTGCTG  
 (1327) ACCCTGACCGTGCAGGCCCGCCAGCTGCTG  
 (1327) ACCCTGACCGTGCAGGCCCGCCAGCTGCTG  
 (1381) ACCCTGACCGTGCAGGCCCGCCAGCTGCTG

1411 1440  
 (1381) AGCGGCATCGTGCAGCAGCAGAACAACTG  
 (1411) AGCGGCATCGTGCAGCAGCAGAACAACTG  
 (1357) AGCGGCATCGTGCAGCAGCAGAACAACTG  
 (1381) AGCGGCATCGTGCAGCAGCAGAACAACTG  
 (1381) AGCGGCATCGTGCAGCAGCAGAACAACTG  
 (1369) AGCGGCATCGTGCAGCAGCAGAACAACTG  
 (1357) AGCGGCATCGTGCAGCAGCAGAACAACTG  
 (1357) AGCGGCATCGTGCAGCAGCAGAACAACTG  
 (1411) AGCGGCATCGTGCAGCAGCAGAACAACTG

1441 1470  
 (1411) CTGGCGGCATCGAGGCCCGCAGCAGCCTG  
 (1441) CTGGCGGCATCGAGGCCCGCAGCAGCCTG  
 (1387) CTGGCGGCATCGAGGCCCGCAGCAGCCTG  
 (1411) CTGGCGGCATCGAGGCCCGCAGCAGCCTG  
 (1411) CTGGCGGCATCGAGGCCCGCAGCAGCCTG  
 (1399) CTGGCGGCATCGAGGCCCGCAGCAGCCTG  
 (1387) CTGGCGGCATCGAGGCCCGCAGCAGCCTG  
 (1387) CTGGCGGCATCGAGGCCCGCAGCAGCCTG  
 (1441) CTGGCGGCATCGAGGCCCGCAGCAGCCTG

1471 1500  
 (1441) CTGCAGCTGACCGTGTGGGGCATCAAGCAG  
 (1471) CTGCAGCTGACCGTGTGGGGCATCAAGCAG  
 (1417) CTGCAGCTGACCGTGTGGGGCATCAAGCAG  
 (1441) CTGCAGCTGACCGTGTGGGGCATCAAGCAG  
 (1441) CTGCAGCTGACCGTGTGGGGCATCAAGCAG  
 (1429) CTGCAGCTGACCGTGTGGGGCATCAAGCAG  
 (1417) CTGCAGCTGACCGTGTGGGGCATCAAGCAG  
 (1417) CTGCAGCTGACCGTGTGGGGCATCAAGCAG  
 (1471) CTGCAGCTGACCGTGTGGGGCATCAAGCAG

1501 1530  
 (1471) CTGCAGGCGCGCGTGTGGCGCTGGAGCGC  
 (1501) CTGCAGGCGCGCGTGTGGCGCTGGAGCGC  
 (1447) CTGCAGGCGCGCGTGTGGCGCTGGAGCGC  
 (1471) CTGCAGGCGCGCGTGTGGCGCTGGAGCGC  
 (1471) CTGCAGGCGCGCGTGTGGCGCTGGAGCGC  
 (1459) CTGCAGGCGCGCGTGTGGCGCTGGAGCGC  
 (1447) CTGCAGGCGCGCGTGTGGCGCTGGAGCGC  
 (1447) CTGCAGGCGCGCGTGTGGCGCTGGAGCGC  
 (1501) CTGCAGGCGCGCGTGTGGCGCTGGAGCGC

1531 1560  
 (1501) TACCTGAAGGACCGAGCTCTGGGCGATC  
 (1531) TACCTGAAGGACCGAGCTCTGGGCGATC

Vall120-Thr202-Ile424-Ala433	(1477)	TACCTGAGGACGACGAGCTGCGGGGATC
Leul22-Ser199-Arg426-Lys432	(1501)	TACCTGAGGACGACGAGCTGCGGGGATC
Leul22-Ser199-Arg426-Gly431	(1501)	TACCTGAGGACGACGAGCTGCGGGGATC
Lys121-Val200-Asn425-Lys432	(1489)	TACCTGAGGACGACGAGCTGCGGGGATC
Vall120-Ile201-Ile424-Ala433	(1477)	TACCTGAGGACGACGAGCTGCGGGGATC
Vall120-Ile201B-Ile424-Ala433	(1477)	TACCTGAGGACGACGAGCTGCGGGGATC
Consensus	(1531)	TACCTGAGGACGACGAGCTGCGGGGATC
Leul22-Ser199 Tryp427-Gly431	(1531)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Vall127-Asn195-Arg426-Gly431	(1561)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Vall120-Thr202-Ile424-Ala433	(1507)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Leul22-Ser199-Arg426-Lys432	(1531)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Leul22-Ser199-Arg426-Gly431	(1531)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Lys121-Val200-Asn425-Lys432	(1519)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Vall120-Ile201-Ile424-Ala433	(1507)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Vall120-Ile201B-Ile424-Ala433	(1507)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Consensus	(1561)	TGGGGCTGCAGCGGCAAGCTGATCTGCACC
Leul22-Ser199 Tryp427-Gly431	(1561)	ACCGCCCTGCCCTGGAACGCGAGCTGGAGC
Vall127-Asn195-Arg426-Gly431	(1591)	ACCGCCCTGCCCTGGAACGCGAGCTGGAGC
Vall120-Thr202-Ile424-Ala433	(1537)	ACCGCCCTGCCCTGGAACGCGAGCTGGAGC
Leul22-Ser199-Arg426-Lys432	(1561)	ACCGCCCTGCCCTGGAACGCGAGCTGGAGC
Leul22-Ser199-Arg426-Gly431	(1561)	ACCGCCCTGCCCTGGAACGCGAGCTGGAGC
Lys121-Val200-Asn425-Lys432	(1549)	ACCGCCCTGCCCTGGAACGCGAGCTGGAGC
Vall120-Ile201-Ile424-Ala433	(1537)	ACCGCCCTGCCCTGGAACGCGAGCTGGAGC
Vall120-Ile201B-Ile424-Ala433	(1537)	ACCGCCCTGCCCTGGAACGCGAGCTGGAGC
Consensus	(1591)	ACCGCCCTGCCCTGGAACGCGAGCTGGAGC
Leul22-Ser199 Tryp427-Gly431	(1591)	AACAAAGGCTTGGACCAAGCTTGGAAACAAC
Vall127-Asn195-Arg426-Gly431	(1621)	AACAAAGGCTTGGACCAAGCTTGGAAACAAC
Vall120-Thr202-Ile424-Ala433	(1567)	AACAAAGGCTTGGACCAAGCTTGGAAACAAC
Leul22-Ser199-Arg426-Lys432	(1591)	AACAAAGGCTTGGACCAAGCTTGGAAACAAC
Leul22-Ser199-Arg426-Gly431	(1591)	AACAAAGGCTTGGACCAAGCTTGGAAACAAC
Lys121-Val200-Asn425-Lys432	(1579)	AACAAAGGCTTGGACCAAGCTTGGAAACAAC
Vall120-Ile201-Ile424-Ala433	(1567)	AACAAAGGCTTGGACCAAGCTTGGAAACAAC
Vall120-Ile201B-Ile424-Ala433	(1567)	AACAAAGGCTTGGACCAAGCTTGGAAACAAC
Consensus	(1621)	AACAAAGGCTTGGACCAAGCTTGGAAACAAC
Leul22-Ser199 Tryp427-Gly431	(1621)	ATGACCTGGATGGAGTGGAGCGCGAGATC
Vall127-Asn195-Arg426-Gly431	(1651)	ATGACCTGGATGGAGTGGAGCGCGAGATC
Vall120-Thr202-Ile424-Ala433	(1597)	ATGACCTGGATGGAGTGGAGCGCGAGATC
Leul22-Ser199-Arg426-Lys432	(1621)	ATGACCTGGATGGAGTGGAGCGCGAGATC
Leul22-Ser199-Arg426-Gly431	(1621)	ATGACCTGGATGGAGTGGAGCGCGAGATC
Lys121-Val200-Asn425-Lys432	(1609)	ATGACCTGGATGGAGTGGAGCGCGAGATC
Vall120-Ile201-Ile424-Ala433	(1597)	ATGACCTGGATGGAGTGGAGCGCGAGATC
Vall120-Ile201B-Ile424-Ala433	(1597)	ATGACCTGGATGGAGTGGAGCGCGAGATC
Consensus	(1651)	ATGACCTGGATGGAGTGGAGCGCGAGATC
Leul22-Ser199 Tryp427-Gly431	(1651)	GACAACTACACCAACCTGATCTACACCGTG
Vall127-Asn195-Arg426-Gly431	(1681)	GACAACTACACCAACCTGATCTACACCGTG
Vall120-Thr202-Ile424-Ala433	(1627)	GACAACTACACCAACCTGATCTACACCGTG
Leul22-Ser199-Arg426-Lys432	(1651)	GACAACTACACCAACCTGATCTACACCGTG
Leul22-Ser199-Arg426-Gly431	(1651)	GACAACTACACCAACCTGATCTACACCGTG
Lys121-Val200-Asn425-Lys432	(1639)	GACAACTACACCAACCTGATCTACACCGTG
Vall120-Ile201-Ile424-Ala433	(1627)	GACAACTACACCAACCTGATCTACACCGTG
Vall120-Ile201B-Ile424-Ala433	(1627)	GACAACTACACCAACCTGATCTACACCGTG
Consensus	(1681)	GACAACTACACCAACCTGATCTACACCGTG

FIG. 5J

Leu122-Ser199 Tryp427-Gly431	(1681)	1711	1740
Val1127-Asn195-Arg426-Gly431	(1711)	ATCGAGGAGAGCCAGAACCCAGCAGGAGAG	
Val1120-Thr202-Ile424-Ala433	(1657)	ATCGAGGAGAGCCAGAACCCAGCAGGAGAG	
Leu122-Ser199-Arg426-Lys432	(1681)	ATCGAGGAGAGCCAGAACCCAGCAGGAGAG	
Leu122-Ser199-Arg426-Gly431	(1681)	ATCGAGGAGAGCCAGAACCCAGCAGGAGAG	
Lys121-Val200-Asn425-Lys432	(1669)	ATCGAGGAGAGCCAGAACCCAGCAGGAGAG	
Val1120-Ile201-Ile424-Ala433	(1657)	ATCGAGGAGAGCCAGAACCCAGCAGGAGAG	
Val1120-Ile201B-Ile424-Ala433	(1657)	ATCGAGGAGAGCCAGAACCCAGCAGGAGAG	
Consensus	(1711)	ATCGAGGAGAGCCAGAACCCAGCAGGAGAG	
Leu122-Ser199 Tryp427-Gly431	(1711)	1741	1770
Val1127-Asn195-Arg426-Gly431	(1741)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Val1120-Thr202-Ile424-Ala433	(1687)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Leu122-Ser199-Arg426-Lys432	(1711)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Leu122-Ser199-Arg426-Gly431	(1711)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Lys121-Val200-Asn425-Lys432	(1699)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Val1120-Ile201-Ile424-Ala433	(1687)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Val1120-Ile201B-Ile424-Ala433	(1687)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Consensus	(1741)	AACGAGCAGGAGCTGCTGGAGCTGGACAAG	
Leu122-Ser199 Tryp427-Gly431	(1741)	1771	1800
Val1127-Asn195-Arg426-Gly431	(1771)	TGGGCCAGCCTGTGGAACCTGGTTCGACATC	
Val1120-Thr202-Ile424-Ala433	(1717)	TGGGCCAGCCTGTGGAACCTGGTTCGACATC	
Leu122-Ser199-Arg426-Lys432	(1741)	TGGGCCAGCCTGTGGAACCTGGTTCGACATC	
Leu122-Ser199-Arg426-Gly431	(1741)	TGGGCCAGCCTGTGGAACCTGGTTCGACATC	
Lys121-Val200-Asn425-Lys432	(1729)	TGGGCCAGCCTGTGGAACCTGGTTCGACATC	
Val1120-Ile201-Ile424-Ala433	(1717)	TGGGCCAGCCTGTGGAACCTGGTTCGACATC	
Val1120-Ile201B-Ile424-Ala433	(1717)	TGGGCCAGCCTGTGGAACCTGGTTCGACATC	
Consensus	(1771)	TGGGCCAGCCTGTGGAACCTGGTTCGACATC	
Leu122-Ser199 Tryp427-Gly431	(1771)	1801	1830
Val1127-Asn195-Arg426-Gly431	(1801)	AGCAAGTGGCTGTGGTACATCAAGATCTTC	
Val1120-Thr202-Ile424-Ala433	(1747)	AGCAAGTGGCTGTGGTACATCAAGATCTTC	
Leu122-Ser199-Arg426-Lys432	(1771)	AGCAAGTGGCTGTGGTACATCAAGATCTTC	
Leu122-Ser199-Arg426-Gly431	(1771)	AGCAAGTGGCTGTGGTACATCAAGATCTTC	
Lys121-Val200-Asn425-Lys432	(1759)	AGCAAGTGGCTGTGGTACATCAAGATCTTC	
Val1120-Ile201-Ile424-Ala433	(1747)	AGCAAGTGGCTGTGGTACATCAAGATCTTC	
Val1120-Ile201B-Ile424-Ala433	(1747)	AGCAAGTGGCTGTGGTACATCAAGATCTTC	
Consensus	(1801)	AGCAAGTGGCTGTGGTACATCAAGATCTTC	
Leu122-Ser199 Tryp427-Gly431	(1801)	1831	1860
Val1127-Asn195-Arg426-Gly431	(1831)	ATCATGATCGTGGGCGGCCCTGGTGGGCGTG	
Val1120-Thr202-Ile424-Ala433	(1777)	ATCATGATCGTGGGCGGCCCTGGTGGGCGTG	
Leu122-Ser199-Arg426-Lys432	(1801)	ATCATGATCGTGGGCGGCCCTGGTGGGCGTG	
Leu122-Ser199-Arg426-Gly431	(1801)	ATCATGATCGTGGGCGGCCCTGGTGGGCGTG	
Lys121-Val200-Asn425-Lys432	(1789)	ATCATGATCGTGGGCGGCCCTGGTGGGCGTG	
Val1120-Ile201-Ile424-Ala433	(1777)	ATCATGATCGTGGGCGGCCCTGGTGGGCGTG	
Val1120-Ile201B-Ile424-Ala433	(1777)	ATCATGATCGTGGGCGGCCCTGGTGGGCGTG	
Consensus	(1831)	ATCATGATCGTGGGCGGCCCTGGTGGGCGTG	
Leu122-Ser199 Tryp427-Gly431	(1831)	1861	1890
Val1127-Asn195-Arg426-Gly431	(1861)	CAGATCGTGTACCCGCTGCTGAGCATCGTG	
Val1120-Thr202-Ile424-Ala433	(1807)	CAGATCGTGTACCCGCTGCTGAGCATCGTG	
Leu122-Ser199-Arg426-Lys432	(1831)	CAGATCGTGTACCCGCTGCTGAGCATCGTG	
Leu122-Ser199-Arg426-Gly431	(1831)	CAGATCGTGTACCCGCTGCTGAGCATCGTG	
Lys121-Val200-Asn425-Lys432	(1819)	CAGATCGTGTACCCGCTGCTGAGCATCGTG	

FIG. 5K

Vall120-Ile201-Ile424-Ala433	(1807)	CGCATCGTGTTCACCGTGCTGAGCATCGTG
Vall120-Ile201B-Ile424-Ala433	(1807)	CGCATCGTGTTCACCGTGCTGAGCATCGTG
Consensus	(1861)	CGCATCGTGTTCACCGTGCTGAGCATCGTG
Leu122-Ser199-Tryp427-Gly431	(1861)	1891
Vall127-Asn195-Arg426-Gly431	(1891)	1920
Vall120-Thr202-Ile424-Ala433	(1891)	AACCGCGTGGCCAGGGCTACAGCCCCCTG
Leu122-Ser199-Arg426-Lys432	(1837)	AACCGCGTGGCCAGGGCTACAGCCCCCTG
Leu122-Ser199-Arg426-Gly431	(1861)	AACCGCGTGGCCAGGGCTACAGCCCCCTG
Lys121-Val200-Asn425-Lys432	(1849)	AACCGCGTGGCCAGGGCTACAGCCCCCTG
Vall120-Ile201-Ile424-Ala433	(1837)	AACCGCGTGGCCAGGGCTACAGCCCCCTG
Vall120-Ile201B-Ile424-Ala433	(1837)	AACCGCGTGGCCAGGGCTACAGCCCCCTG
Consensus	(1891)	AACCGCGTGGCCAGGGCTACAGCCCCCTG
Leu122-Ser199-Tryp427-Gly431	(1891)	1921
Vall127-Asn195-Arg426-Gly431	(1921)	1950
Vall120-Thr202-Ile424-Ala433	(1867)	AGCTTCCAGACCCGCTTCCCGGCCCGCCG
Leu122-Ser199-Arg426-Lys432	(1891)	AGCTTCCAGACCCGCTTCCCGGCCCGCCG
Leu122-Ser199-Arg426-Gly431	(1891)	AGCTTCCAGACCCGCTTCCCGGCCCGCCG
Lys121-Val200-Asn425-Lys432	(1879)	AGCTTCCAGACCCGCTTCCCGGCCCGCCG
Vall120-Ile201-Ile424-Ala433	(1867)	AGCTTCCAGACCCGCTTCCCGGCCCGCCG
Vall120-Ile201B-Ile424-Ala433	(1867)	AGCTTCCAGACCCGCTTCCCGGCCCGCCG
Consensus	(1921)	AGCTTCCAGACCCGCTTCCCGGCCCGCCG
Leu122-Ser199-Tryp427-Gly431	(1921)	1980
Vall127-Asn195-Arg426-Gly431	(1951)	GGCCCGGAGCCCGGAGGSCATCGAGGAG
Vall120-Thr202-Ile424-Ala433	(1897)	GGCCCGGAGCCCGGAGGSCATCGAGGAG
Leu122-Ser199-Arg426-Lys432	(1921)	GGCCCGGAGCCCGGAGGSCATCGAGGAG
Leu122-Ser199-Arg426-Gly431	(1921)	GGCCCGGAGCCCGGAGGSCATCGAGGAG
Lys121-Val200-Asn425-Lys432	(1909)	GGCCCGGAGCCCGGAGGSCATCGAGGAG
Vall120-Ile201-Ile424-Ala433	(1897)	GGCCCGGAGCCCGGAGGSCATCGAGGAG
Vall120-Ile201B-Ile424-Ala433	(1897)	GGCCCGGAGCCCGGAGGSCATCGAGGAG
Consensus	(1951)	GGCCCGGAGCCCGGAGGSCATCGAGGAG
Leu122-Ser199-Tryp427-Gly431	(1951)	2010
Vall127-Asn195-Arg426-Gly431	(1981)	GAGGGCGGCGAGCGGAGCGGACCGGAGC
Vall120-Thr202-Ile424-Ala433	(1927)	GAGGGCGGCGAGCGGAGCGGACCGGAGC
Leu122-Ser199-Arg426-Lys432	(1951)	GAGGGCGGCGAGCGGAGCGGACCGGAGC
Leu122-Ser199-Arg426-Gly431	(1951)	GAGGGCGGCGAGCGGAGCGGACCGGAGC
Lys121-Val200-Asn425-Lys432	(1939)	GAGGGCGGCGAGCGGAGCGGACCGGAGC
Vall120-Ile201-Ile424-Ala433	(1927)	GAGGGCGGCGAGCGGAGCGGACCGGAGC
Vall120-Ile201B-Ile424-Ala433	(1927)	GAGGGCGGCGAGCGGAGCGGACCGGAGC
Consensus	(1981)	GAGGGCGGCGAGCGGAGCGGACCGGAGC
Leu122-Ser199-Tryp427-Gly431	(1981)	2040
Vall127-Asn195-Arg426-Gly431	(2011)	AGCCCGCTGGTGACAGGCGCTGCTGGCCCTG
Vall120-Thr202-Ile424-Ala433	(1957)	AGCCCGCTGGTGACAGGCGCTGCTGGCCCTG
Leu122-Ser199-Arg426-Lys432	(1981)	AGCCCGCTGGTGACAGGCGCTGCTGGCCCTG
Leu122-Ser199-Arg426-Gly431	(1981)	AGCCCGCTGGTGACAGGCGCTGCTGGCCCTG
Lys121-Val200-Asn425-Lys432	(1969)	AGCCCGCTGGTGACAGGCGCTGCTGGCCCTG
Vall120-Ile201-Ile424-Ala433	(1957)	AGCCCGCTGGTGACAGGCGCTGCTGGCCCTG
Vall120-Ile201B-Ile424-Ala433	(1957)	AGCCCGCTGGTGACAGGCGCTGCTGGCCCTG
Consensus	(2011)	AGCCCGCTGGTGACAGGCGCTGCTGGCCCTG
Leu122-Ser199-Tryp427-Gly431	(2011)	2070
Vall127-Asn195-Arg426-Gly431	(2041)	ATCTGGGAGGAGCTGCGAGGCTGCTGGCTG
Vall120-Thr202-Ile424-Ala433	(1987)	ATCTGGGAGGAGCTGCGAGGCTGCTGGCTG

FIG. 5L

Leu122-Ser199-Arg426-Lys432	(2011)	ATCTGGGAGGACCTGGCGAGCCTGTGCGCTG
Leu122-Ser199-Arg426-Gly431	(2011)	ATCTGGGAGGACCTGGCGAGCCTGTGCGCTG
Lys121-Val200-Asn425-Lys432	(1999)	ATCTGGGAGGACCTGGCGAGCCTGTGCGCTG
Val120-Ile201-Ile424-Ala433	(1987)	ATCTGGGAGGACCTGGCGAGCCTGTGCGCTG
Val120-Ile201B-Ile424-Ala433	(1987)	ATCTGGGAGGACCTGGCGAGCCTGTGCGCTG
Consensus	(2041)	ATCTGGGAGGACCTGGCGAGCCTGTGCGCTG
Leu122-Ser199-Tryp427-Gly431	(2041)	TTGAGCTACCACCGCCTGGCGAGCTGATC
Val127-Asn195-Arg426-Gly431	(2071)	TTGAGCTACCACCGCCTGGCGAGCTGATC
Val120-Thr202-Ile424-Ala433	(2017)	TTGAGCTACCACCGCCTGGCGAGCTGATC
Leu122-Ser199-Arg426-Lys432	(2041)	TTGAGCTACCACCGCCTGGCGAGCTGATC
Leu122-Ser199-Arg426-Gly431	(2041)	TTGAGCTACCACCGCCTGGCGAGCTGATC
Lys121-Val200-Asn425-Lys432	(2029)	TTGAGCTACCACCGCCTGGCGAGCTGATC
Val120-Ile201-Ile424-Ala433	(2017)	TTGAGCTACCACCGCCTGGCGAGCTGATC
Val120-Ile201B-Ile424-Ala433	(2017)	TTGAGCTACCACCGCCTGGCGAGCTGATC
Consensus	(2071)	TTGAGCTACCACCGCCTGGCGAGCTGATC
Leu122-Ser199-Tryp427-Gly431	(2071)	CTGATCGCGCGCCGCATCGTGAGCTGCTG
Val127-Asn195-Arg426-Gly431	(2101)	CTGATCGCGCGCCGCATCGTGAGCTGCTG
Val120-Thr202-Ile424-Ala433	(2047)	CTGATCGCGCGCCGCATCGTGAGCTGCTG
Leu122-Ser199-Arg426-Lys432	(2071)	CTGATCGCGCGCCGCATCGTGAGCTGCTG
Leu122-Ser199-Arg426-Gly431	(2071)	CTGATCGCGCGCCGCATCGTGAGCTGCTG
Lys121-Val200-Asn425-Lys432	(2059)	CTGATCGCGCGCCGCATCGTGAGCTGCTG
Lys121-Val200-Ile424-Ala433	(2047)	CTGATCGCGCGCCGCATCGTGAGCTGCTG
Val120-Ile201B-Ile424-Ala433	(2047)	CTGATCGCGCGCCGCATCGTGAGCTGCTG
Consensus	(2101)	CTGATCGCGCGCCGCATCGTGAGCTGCTG
Leu122-Ser199-Tryp427-Gly431	(2101)	GGCGCGCGCGCTGGGAGGCCCTGAAGTAC
Val127-Asn195-Arg426-Gly431	(2131)	GGCGCGCGCGCTGGGAGGCCCTGAAGTAC
Val120-Thr202-Ile424-Ala433	(2077)	GGCGCGCGCGCTGGGAGGCCCTGAAGTAC
Leu122-Ser199-Arg426-Lys432	(2101)	GGCGCGCGCGCTGGGAGGCCCTGAAGTAC
Leu122-Ser199-Arg426-Gly431	(2101)	GGCGCGCGCGCTGGGAGGCCCTGAAGTAC
Lys121-Val200-Asn425-Lys432	(2089)	GGCGCGCGCGCTGGGAGGCCCTGAAGTAC
Val120-Ile201-Ile424-Ala433	(2077)	GGCGCGCGCGCTGGGAGGCCCTGAAGTAC
Val120-Ile201B-Ile424-Ala433	(2077)	GGCGCGCGCGCTGGGAGGCCCTGAAGTAC
Consensus	(2131)	GGCGCGCGCGCTGGGAGGCCCTGAAGTAC
Leu122-Ser199-Tryp427-Gly431	(2131)	TGGGGCACTTCTGCGAGTACTGGATCCAG
Val127-Asn195-Arg426-Gly431	(2161)	TGGGGCACTTCTGCGAGTACTGGATCCAG
Val120-Thr202-Ile424-Ala433	(2107)	TGGGGCACTTCTGCGAGTACTGGATCCAG
Leu122-Ser199-Arg426-Lys432	(2131)	TGGGGCACTTCTGCGAGTACTGGATCCAG
Leu122-Ser199-Arg426-Gly431	(2131)	TGGGGCACTTCTGCGAGTACTGGATCCAG
Lys121-Val200-Asn425-Lys432	(2119)	TGGGGCACTTCTGCGAGTACTGGATCCAG
Val120-Ile201-Ile424-Ala433	(2107)	TGGGGCACTTCTGCGAGTACTGGATCCAG
Val120-Ile201B-Ile424-Ala433	(2107)	TGGGGCACTTCTGCGAGTACTGGATCCAG
Consensus	(2161)	TGGGGCACTTCTGCGAGTACTGGATCCAG
Leu122-Ser199-Tryp427-Gly431	(2161)	GAGCTGAAGAACAGGCCCTGAGGCTGTTC
Val127-Asn195-Arg426-Gly431	(2191)	GAGCTGAAGAACAGGCCCTGAGGCTGTTC
Val120-Thr202-Ile424-Ala433	(2137)	GAGCTGAAGAACAGGCCCTGAGGCTGTTC
Leu122-Ser199-Arg426-Lys432	(2161)	GAGCTGAAGAACAGGCCCTGAGGCTGTTC
Leu122-Ser199-Arg426-Gly431	(2161)	GAGCTGAAGAACAGGCCCTGAGGCTGTTC
Lys121-Val200-Asn425-Lys432	(2149)	GAGCTGAAGAACAGGCCCTGAGGCTGTTC
Val120-Ile201-Ile424-Ala433	(2137)	GAGCTGAAGAACAGGCCCTGAGGCTGTTC
Val120-Ile201B-Ile424-Ala433	(2137)	GAGCTGAAGAACAGGCCCTGAGGCTGTTC
Consensus	(2191)	GAGCTGAAGAACAGGCCCTGAGGCTGTTC

2221

2250

FIG. 5M

Leu122-Ser199 Tryp427-Gly431	(2191)	GACGCCATCGCCATCGCCGTGGCCGAGGGC
Val1127-Asn195-Arg426-Gly431	(2221)	GACGCCATCGCCATCGCCGTGGCCGAGGGC
Val1120-Thr202-Ile424-Ala433	(2167)	GACGCCATCGCCATCGCCGTGGCCGAGGGC
Leu122-Ser199-Arg426-Lys432	(2191)	GACGCCATCGCCATCGCCGTGGCCGAGGGC
Leu122-Ser199-Arg426-Gly431	(2191)	GACGCCATCGCCATCGCCGTGGCCGAGGGC
Lys121-Val200-Asn425-Lys432	(2179)	GACGCCATCGCCATCGCCGTGGCCGAGGGC
Val1120-Ile201-Ile424-Ala433	(2167)	GACGCCATCGCCATCGCCGTGGCCGAGGGC
Val1120-Ile201B-Ile424-Ala433	(2167)	GACGCCATCGCCATCGCCGTGGCCGAGGGC
Consensus	(2221)	GACGCCATCGCCATCGCCGTGGCCGAGGGC
Leu122-Ser199 Tryp427-Gly431	(2221)	ACCGACCGCATCATCGAGGTGGCCGAGGCG
Val1127-Asn195-Arg426-Gly431	(2251)	ACCGACCGCATCATCGAGGTGGCCGAGGCG
Val1120-Thr202-Ile424-Ala433	(2197)	ACCGACCGCATCATCGAGGTGGCCGAGGCG
Leu122-Ser199-Arg426-Lys432	(2221)	ACCGACCGCATCATCGAGGTGGCCGAGGCG
Leu122-Ser199-Arg426-Gly431	(2221)	ACCGACCGCATCATCGAGGTGGCCGAGGCG
Lys121-Val200-Asn425-Lys432	(2209)	ACCGACCGCATCATCGAGGTGGCCGAGGCG
Val1120-Ile201-Ile424-Ala433	(2197)	ACCGACCGCATCATCGAGGTGGCCGAGGCG
Val1120-Ile201B-Ile424-Ala433	(2197)	ACCGACCGCATCATCGAGGTGGCCGAGGCG
Consensus	(2251)	ACCGACCGCATCATCGAGGTGGCCGAGGCG
Leu122-Ser199 Tryp427-Gly431	(2251)	ATCGGCCGCGCCTTCCTGCACATCCCCGC
Val1127-Asn195-Arg426-Gly431	(2281)	ATCGGCCGCGCCTTCCTGCACATCCCCGC
Val1120-Thr202-Ile424-Ala433	(2227)	ATCGGCCGCGCCTTCCTGCACATCCCCGC
Leu122-Ser199-Arg426-Lys432	(2251)	ATCGGCCGCGCCTTCCTGCACATCCCCGC
Leu122-Ser199-Arg426-Gly431	(2251)	ATCGGCCGCGCCTTCCTGCACATCCCCGC
Lys121-Val200-Asn425-Lys432	(2239)	ATCGGCCGCGCCTTCCTGCACATCCCCGC
Val1120-Ile201-Ile424-Ala433	(2227)	ATCGGCCGCGCCTTCCTGCACATCCCCGC
Val1120-Ile201B-Ile424-Ala433	(2227)	ATCGGCCGCGCCTTCCTGCACATCCCCGC
Consensus	(2281)	ATCGGCCGCGCCTTCCTGCACATCCCCGC
Leu122-Ser199 Tryp427-Gly431	(2281)	CGCATCCGCGAGGGCTTCGAGCGCGCCCTG
Val1127-Asn195-Arg426-Gly431	(2311)	CGCATCCGCGAGGGCTTCGAGCGCGCCCTG
Val1120-Thr202-Ile424-Ala433	(2257)	CGCATCCGCGAGGGCTTCGAGCGCGCCCTG
Leu122-Ser199-Arg426-Lys432	(2281)	CGCATCCGCGAGGGCTTCGAGCGCGCCCTG
Leu122-Ser199-Arg426-Gly431	(2281)	CGCATCCGCGAGGGCTTCGAGCGCGCCCTG
Lys121-Val200-Asn425-Lys432	(2269)	CGCATCCGCGAGGGCTTCGAGCGCGCCCTG
Val1120-Ile201-Ile424-Ala433	(2257)	CGCATCCGCGAGGGCTTCGAGCGCGCCCTG
Val1120-Ile201B-Ile424-Ala433	(2257)	CGCATCCGCGAGGGCTTCGAGCGCGCCCTG
Consensus	(2311)	CGCATCCGCGAGGGCTTCGAGCGCGCCCTG
Leu122-Ser199 Tryp427-Gly431	(2311)	CTGTAACTCGAG
Val1127-Asn195-Arg426-Gly431	(2341)	CTGTAACTCGAG
Val1120-Thr202-Ile424-Ala433	(2287)	CTGTAACTCGAG
Leu122-Ser199-Arg426-Lys432	(2311)	CTGTAACTCGAG
Leu122-Ser199-Arg426-Gly431	(2311)	CTGTAACTCGAG
Lys121-Val200-Asn425-Lys432	(2299)	CTGTAACTCGAG
Val1120-Ile201-Ile424-Ala433	(2287)	CTGTAACTCGAG
Val1120-Ile201B-Ile424-Ala433	(2287)	CTGTAACTCGAG
Consensus	(2341)	CTGTAACTCGAG

FIG. 5N



## SEQ ID NO:3 VAL120-ALA204

GAATTCGCCACCATGGATGCAATGAAGAGAGGGGCTGCTGCTGTGCTGCTGCTGTGTGGAGCA  
GTCTTCGTTTCGCCCAAGCGCGTGGAGAAAGCTGTGGGTGACCGTGACTACGGCGTGCCCGTG  
TGGAAGGAGGGCCACCCACCCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGCACACCGAGGT  
GCACAACGTGTGGGCCACCCACGCTGCGTGCCCAACCGCAACCCCAAGGAGATCTGTGCT  
GGAGAAACGTGACCGGAGAATTCACACTGTGGAAAGAAACAATGTTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAAGAGCCTGAAGCCCTGCGTGGCGCGCGCGCTGCCCAA  
GGTGAGCTTCGAGCCCATCCCACTCACTGCGCCCGCGCGCGCTTCGCCATCTCTGAAAGT  
CAACGCACAAGAGTTCAACGGCAAGCGCCCTGCACCAACGTGAGCACCGTGACGTGCACCC  
ACGGCATCCGCCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGC  
GTGGTGATCCGCAGCGAGAATTCACCGACAACGCCAAGACCATCATCGTGACGTGAAGGA  
GAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCAACATCGGCC  
CCGGCCGCGCTTCTACGCCACCGCGCATATCGGCGACATCCGCCAGGCCACTGCAACA  
TCAGGGCGGAAGGTGGAACAACACCTGAAAGCAGATCGTGACCAAGCTGCAGGCCAGTTTC  
GGCAACAAGACCATCTGTGTTCAAGCAGAGCAGCGCGCGGCAACCCGAGATCTGTGATGCACAG  
CTTCAACTGCGGGCGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGACCTGGAA  
CAACACCATCGGCCCCAACACACCAACGGCACCATCACTGCTGCCCTGCGCATCAAGCAGA  
TCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCATCCGCGGCGAGATC  
CGCTGCAGCAGCAACATCAACCGCGCTGCTGCTGACCCGCGACGGCGGCAAGGAGATCAGCAA  
CACCACCGAGATCTTCCGCCCGCGCGCGCGGACATGCGCGACAACCTGCGCAGCGAGCTGT  
ACAAAGTACAAGGTGGTGAAGATCGAGGCCCTTGGGCGTGGCCCCCAACCAAGGCCAAGCGCGC  
GTGGTGACGCGGAGAAAGCGCGCGGTGACCTGGGCGCCATGTTCTGGGCTTCTGGGCGCC  
GCGCGCAGCACCATGGGCGCCCGCAGCTGACCTGACCGTGACAGGCCCGCCAGCTGCTGAG  
CGGCATCGTGCAGCAGCAGAACAACTGCTGGCGGCCATCGAGGCCAGCAGCACCTGCTGC  
AGCTGACCGGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGTGGCCGTGGAGCGCTACCTG  
AAGGACCAAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACACCGCGT  
GCCCTGGAAACGCCAGCTGGAGCAACAAGAGCCTGGACCAAGATCTGGAAACAATGACCTGGA  
TGGAGTGGGAGCGCGAGATCGACAACCTACCAACCTGATCTACACCTGATCGAGGAGAGC  
CAGAACCAAGAGGAGAAAGAACGAGCAGGAGTGTCTGGAGCTGGACAAGTGGGCCAGCCTGT  
GGAACTGGTTCGACATCAGCAAGTGGCTGTGTTACATCAAGATCTTATCATGATCGTGGGG  
GCCTGTGGGCGCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGTGGCGCAGGGCT  
ACAGCCCCCTGAGCTTCCAGACCCGCTTCCCGCCCCCGCGGCCCCGACCGCCCCGAGGGCA  
TCGAGGAGGAGGGCGGCGAGCGCGACCGGACCGCAGCAGCCCCCTGGTGCAAGGCGCTGTG  
GCCCTGATCTGGGACGACCTGCGCAGCCTGTGCTGTTACAGTACCAACCGCTGCGCGACCTG  
ATCTGATTCGCGCGCCGCATCGTGGAGCTGCTGGGCGCGCGGCTGGGAGGCCCTGAAGTAC  
TGGGGCAAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGCTGAGCCTGTTGCA  
CGCCATCGCCATCGCGTGGCGGAGGGCACCGCATCATCGAGGTGGCCAGCGCATCG  
GCGCGCCTTCTGCACATCCCCCGCGCATCGGCCAGGGCTTCGAGCGCGCCTGCTGTAACTC  
TCGAG

FIG. 6

## SEQ ID NO:4 VAL120-ILE201

GAATTTCGCCACCATGGATGCAATGAAGAGAGGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA  
GTCTTCGTTTCGCCACGCGCTGGAGAGAAGCTGTGGGTGACCGTGTAACACGCGCTGCCCGTG  
TGGAAGGAGGCCACCAACCAACCTGTTCTCGCCAGCGACGCGCAAGGCCTACGACAACCGAGGT  
GCACAACGCTGTGGGCCACCAACGCTGCTGCTGCCACCGACCCCAACCCCAAGGAGATCTGTGCT  
GGAGAAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGTGTGGAGCAGATGCACGAG  
GACATCATACGCTGTGTGGGACAGAGCTGGAAGCCCTGCTGTGGGCGGCATCAACCGGCGCT  
CCCCAAGGTGAGCTTCGAGCCCATCCCATCCACTACTGCGCCCGCGCGGCTTCGCCATCTT  
GAAGTGCAACGACAAGAAAGTTCAACGCGACGCGCCCTGCAACCAAGTGAAGCAACCGTGCAAT  
GCACCCACGCGATCCGCCCCGTGTGTGAGCAACCCAGCTGCTGTGAACGCGACGCTGGCCGAG  
GAGGGCGTGTGTATCCGCAGCGAGAACTTCAACGACAACGCCAAGACCATCATCTGTGCAGCT  
GAAGGAGAGCGTGTGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCA  
TCGGCCCCGGCCGCGCTTCTACGCCACCGCGGACATCATCGGCGACATCCGCCAGGCCCACT  
GCAACATCAGCGCGGAGAAGTGAACAACACCTGAAAGCAGATCGTGACCAAGCTGCAGGGCC  
CAGTTTCGGCAACAAGACCATCGTGTTCGAAGCAGAGCAGCGGCGCGACCCCGAGATCGTGAT  
GCACAGCTTCAACTGCGCGCGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCAC  
CTGGAACAACACCATCGGCCCAACAACAACCAACGCGACCATCAACCTGCGCTGCCGATCA  
AGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCATCCGCGGC  
CAGATCCGCTGCAGCAGCAACATCACCGGCTGTGTGTGACCCGCGACGCGGCAAGGAGAT  
CAGCAACACCAACGAGATCTTCCGCCCGGCGGCGGCGACATGCGCGACAACCTGGCGCAGCG  
AGCTGTACAAGTACAAGGTGTGAAGATCGAGGCCCTGGGCGTGGGCCCAACCAAGGCCAAG  
CGCCGCGTGTGTGAGCGCGAGAAGCGCGCGGTGACCTGGGCGGCAATGTTCTGGGCTTCTGT  
GGCGCGCGCGGCGACCATGGGCGCGCGCGAGCTGACCTGACCGTGCAAGGCCCGCGCAGCT  
GCTGAGCGGCATCGTGACGACGAGACAACCTGTGTGCGCGCCATCGAGGCCCGACGACACC  
TGCTGCACTGACCGTGTGGGGCATCAAGCAGCTGCAAGGCCCGCGTGTGTGCGCTGTGAGCGC  
TACCTGAAGGACCAAGCAGCTGTGTGGGCATCTGGGCTGTGACGCGCAAGCTGATGTGCAACAC  
CGCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGAGCCAGATCTGGAACAACATGA  
CCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACCAACCTGATCTACACCTGATCGAG  
GAGAGCCAGAACGACGAGGAGAAGACGAGCAGGAGCTGTGTGAGCTGGACAAGTGGGCCA  
GCCTGTGGAACCTGTTGACATCAGCAAGTGGCTGTGATACATCAAGATCTTATCATGATCG  
TGGGCGCGCTGTGTGGGCTGTGCGATCGTGTTCACCGTGTGAGCATCGTGAACCGCGTGTGCGC  
AGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCGCGCGGCCCGACCGGCCCG  
AGGGCATCGAGGAGGAGGGCGCGGACGCGACCGGACCGCAGCGACCGCCCTGTTGACAGG  
CTGTCTGCCCCTGATCTGGGACGAACTGCGCAGCTGTGCTGTTCAGCTAACACCGGCTCGG  
CGACCTGATCTGATCGCCGCCGATCTGTGAGCTGTGTGGGCGGCCGCGCTGGGAGGCCCT  
GAAATGATCTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAAGCGCGTGTGAGC  
TGTTCGACGCCATCGCCATCGCGTGGCGGAGGGCACCGCAGCTATCATCAGGTTGGCCAGC  
GCATCGGCGCGCTTCTGACATATCCCGCGCGCATCCGCCAGGGCTTCGAGCGCGGCTGCG  
TGTAACCTCAG

FIG. 7

## SEQ ID NO:5 VAL120-ILE201B

GAATTCCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCAGCTCTTGG  
 TTTCGCCACAGCGCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCTGTGGAAAGGAGGCCA  
 CCAACACCTGTGTTCTGCGCCACGGACGCCAAGGCGCTACGACACCGAGGTGCAACAACGTGTGGCCACCC  
 ACGCCTGCGTGCCTCCGACCGCAACCCCAAGGAGATCGTGTGGAGAACGTGACCGAGAACTTCAACA  
 TGTGGAAGAACCAATGTTGGAGCAGATGCAACGAGGACATCATCAGCCTGTGGGACCAAGCGCTGAAGC  
 CTTGCGTGCCTGGCATCACCCAGGCGTGCCTCAAGTTGAGTTCAACGGCAGCGGCCCTGCAACCAAGT  
 CCCCAGCGCTTGGCACTCTGAAGTGCACCGCATCCGCGCCGTGTTGAGCAACCGAGCTGCTGCTGAACGGCAGCT  
 GAGCACCGTGCAGTGCACCCACCGCATCCGCGCCGTGTTGAGCAACCGAGCTGCTGCTGAACGGCAGCT  
 GGCAGAGGAGCGCTGTTGATCCGACAGGAGAACTTCAACGCAACCGCAAGACCATCATCTGTGTCAGCT  
 GAAGGAGAGCGTGGAGATCAACTGCACCGCCCAACAACAACCGCAGAGCATCAACATCGCCGCC  
 CCGCGCGCTTCTACGCCAAGCGGACATCATCGGCGACATCCGCCAGGCCACTGCAACATCAGCGGC  
 GAGAAGTGGAAACAACCGCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAAGTTGCGCAACAAGACCATC  
 GTGTTCAAGCAGAGCAGCGCGCGGACCGGAGATGTTGATGACAGCTTCAACTGCGCGCGGAGTTTC  
 TTCTACTGCAACAGCACCGAGCTGTTCAACAGCAGCTGGAACAACACCATCGGCCCAACAACAACAAC  
 GGCAACATCAACCTGCGCTGCGCATCAAGCAGATCATCAACCGCTGGCAGGAGTGGGCAAGGCCATG  
 TACGCCCGCCCATCGCGCGGAGATCCGTGACGACGAACAATCACCAGCGCTGCTGCTGACCGCGAGC  
 GCGGCAAGGAGATCAGCAACCAACCGAGATCTTCCGCCCGCGCGCGGACATCGCGGCAACTGGC  
 GCAGCGAGCTGTACAAATACAAGGTGTTGAAGATCGAGCCCTGGGCGTGGCCCCCAACCAAGGCCAAGC  
 GCCCGTGTGTGTCAGCGCGGAGAAAGCGCGCGTGAACCTGGGCGCCATGTTCTGGGCTTCTGGGCGCGC  
 CGGCAGCACCATGGGCGCGCGCAGCTGACCTGACCGTGCAGGCGCGCGAGCTGCTGAGCGGCATCGT  
 GCAGCAGCAGAACAACTGCTGCGCGCATCGAAGGCCAGCAGCACTGTGACGCTGACCGTGTGGGG  
 CATCAAGCAGCTGACGAGCGCGCTGCTGGCGGTGGAGCGCTACCTGAAGGACCAAGCACTGCTGGGCAT  
 CTGGGGCTGACGCGGCAAGCTGATCTGCACCAACCGCGTGCCTGGAAAGCGCAGCTGGAGCAACAAGAG  
 CCTGGACCAATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGCAAACTACCAACCT  
 GATCTACACCTGATCGAGGAGAGCCAGAACCAGCAGGAGAAAGAACGAGCAGGAGCTGTGGAGCTGG  
 ACAAGTGGGCGAGCTGTGGAACTGTTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT  
 GATCTGGGCGGCGTGTGGCGCTGCGCATCGTGTTCACCGTGTGAGCATCTGTGAACCGCGTGGCGCAG  
 GGTACTACGCCCTGAGCTTCCAGACCGCTTTCGCCCGCCCGCGCGGCCAGCGCCCGAGCGCATCT  
 AGGAGAGGGCGCGCGAGCGGACCGCGACCGCGACCGCCCTGGTGACCGCGCTGCTGGCGCTGATCT  
 GGGACGACCTGCGCAGCTGTGCTGTTTCAAGTACCAACCGCTGCGGACATGATCTGATCTGATCGCGCGC  
 CATGTGGAGCTGTGGGCGCGCGCGCTGGAGGCCCTGAAGTACTGGGCAACCTGCTGCACTGATCTG  
 GATCCAGGAGCTGAAGAACAGCGCGCTGAGCTGTTCAGCGCATCCCATCGCGCTGGCGGAGGCGAC  
 CGACCGCATCATCAGAGTGGCCAGCGCATCGCGCGCTTCTGTCACATCCCCCGCGCATCCGCCAG  
 GGCTTCGAGCGCGCTGCTGTAACTCGAGCGTGTCT

FIG. 8

## SEQ ID NO:6 LYS121-VAL200

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA  
GTCTTCGTTTCGCCCAACGCGCGTGGAGAAGCTGTGGGTGACCGTGTAACGCGCGTGCCCGTG  
TGGAAAGGAGGCCACCAACCACTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT  
GCACAAACGCTGTGGGCCACCAACGCGCTGCGTGCCACCGACCCCAACCCCAAGGAGATCGTGT  
GGAGAACGTGACCGAGAACTTCAACATGTGGGAAGAACAAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACAGAGCCTGGAAGCCTGCGTGAAGGCCCGCTGATCATCCCA  
GGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCATCCACTACTGCGCCCCCGCGGCTTCGC  
CATCTCTGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCG  
TGCAGTGACCCACGGCATCCGCGCCGCTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGG  
CCGAGGAGGGCGTGGTGATCCGACGCGAGAATTCAACGACAACGCCAAGACCATCATCGTG  
CAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACCCGCAAGAGCAT  
CACCATCGGCCCCGCGCGCTTCTACGCCACCGCGACATCATCGGCGACATCCGCGAGGC  
CCACTGCAACATCAGCGCGGAGAAAGTGGAACAACACCTGAAAGCAGATCGTGACCAAGCTGC  
AGGCCACAGTTCGGCAACAAGACCATCGTGTTCAGCAGAGCAGCGCGCGGACCCCGAGATC  
GTGATGCACAGCTTCAACTGGGGCGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAAC  
AGCACCTGGAACAACACCATCGGCCCAACAACAACCGCACCATCACCTGCGCTGCGG  
CATCAAGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCATCC  
GCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCTGCTGCTGACCCGCGCAGCGCGGCAAG  
GAGATCAGCAACACCCAGAGATCTTCCGCCCGCGCGCGCGACATGCGCGACAACCTGGCG  
CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCCTGGGCGTGGCCCCACCAAGG  
CCAAGCGCGCGTGGTGACGCGGAGAAGCGCGCGCTGACCTGGGCGCATGTTTCTGGGG  
TTCTCTGGGCGCGCGCGCAGCAACATGGGCGCCCGCAGCCTGACCTGACCGTGACAGGCCCGC  
CAGCTGTGTGACGGCATCGTGCAGCAGCAGAAACCTGCTGCGCGCATCGAGGCCGATCAGG  
GCACCTGCTCAGCTGACCGTGTGGGGCATCAAGCAGCTGACGGCCCGCGTGTGGCGGTGG  
AGCGCTACTGAAGGACAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCG  
ACCAACCGCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCCTGGACCATGATCTGAACAA  
CATGACCTGGATGGAAGTGGGAGCGCGAGATCGACAACACTACCAACCTGATCTACACCCCTGA  
TCGAGGAGAGCCAGAACCCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTG  
GGCCACGCTGTGGAACCTGGTTCGACATCAGCAAGTGCTGTGTTACATCAAGATCTTACATCAT  
GATCTGGGGCGCGCTGGTGGGCTTGCATCTGTTCACCGTGCTGAGCATCTGTAACCGCGT  
GCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCGCCCCCGCGCGGCCGACCG  
CCCCGAGGGCATCGAGGAGGAGGGCGCGAGCGCAGCCGACCCGACGACGCCCTGGTG  
ACGGCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCTGTTTCACTACCAACCGCC  
TGCGGCACTGATCTGATCGCGCCCGCATCTGGAGCTGCTGGGCGCGCGCGCTGGGAGG  
CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG  
AGCCTGTTTCAGCCCATCGCCATCGCGTGGCCGAGGGCACCCAGCGCATCATCGAGGTGGCC  
CAGCGCATCGGCCGCGCTTCTGACATCCCGCGCGCATCCGCAAGGCTTCGAGCGCGCC  
CTGCTGTAACCTGAGCGTGTCT

FIG. 9

## SEQ ID NO:7: LEU122-SER199

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA  
GTCTTCGTTTCGCCACGCGCGTGGAGAAAGCTGTGGGTGACCGGTGACTACGGCGTGCCTGTG  
TGGAAAGGAGGCCACCAACCCCTGTCTTGGCCAGCGACGCCAAGGCTACGACACCGAGGT  
GCACAACCGTGTGGGCCACCAACCGCTGCGTGCCACCGACCCCAACCCCAAGGAGATCGTGTGCT  
GGAGAAAGCTGACCGGAGAACTTCAACATGTGGAAAGAAACAATGGTGGAGCAGATGCACGAG  
GACATCATGAGCCTGTGGGACCAAGCCTGAAGCCCTGCGTGAAGCTGGGGCAACAGCGTGAT  
CACCACGCGCTGCCCCAAGGTGAGCTTCGAGCCCATCCCATCCACTGCGGCCCGCGCGG  
CTTCGCCATCTGAAAGTGAACGACAAGAAAGTTCAACGGCAGCGGCCCTGCACCAACGTGA  
GCACCGTGCAGTGACCCACGGCATTCGCGCCGTGGTGAGCACCCAGCTGCTGTCTGAACGGC  
AGCCTGGCCGAGGAGGGCGTGGTGATCCGCGAGCGAAGCTTCACCGACAACGCCAAGACCAT  
CATCGTGACGTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCA  
AGAGCATCAACCATCGGCCCGCGCGCCTTCTACGCCACCGCGACATCATCGGCCACATCC  
GCCAGGCCCATGCAACATCAGCGCGGAGAAAGTGGAAACAACCCCTGAAGCAGATCGTGACC  
AAGCTGCAGGCCCAAGTTTCGGCAACAAGACCATCGTGTTCGAAGCAGAGCAGCGCGCGCACCC  
CGAGATCGTGATGCACAGCTTCAACTGCGCGCGCGAGTTCTTCTACTGCAACAGCACCCAGCT  
GTTCAACAGCACTTGGAAACAACCATCGGCCCAACAACAACCAACGGCACCATCAACCTGTC  
CCTGCGCATCAAGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCC  
CCCATCCGCGCGCAGATCCGCTGCAGCAGCAACATCACCGGCTGCTGCTGACCCGCGACGCG  
GGCAAGGAGATGACCAACACCCAGAGATCTTCGCGCCCGCGCGCGCGACATGGCGGCCAA  
CTGGCGCAGCTGATCAAGTACAAGGTGGTGAAGATCGAGCCCTGGGCTGCGCGCGCA  
CCAAGGCCCAAGCGCCCGTGGTGACGCGGAGAAAGCGCGCGCTGACCTGGGCGCCATGTTTC  
TGGGCTTCTGGGCGCGCGCGGACGACCATGGGCGCCCGACGCTGACCTGACCGTGCA  
GCCCCGACGCTGCTGAGCGGCATCGTGACGACGAGAAACAACCTGCTGCGCGCATCGAGGC  
CCAGCAGCACTGCTGACAGTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGTGG  
CCGTGGAGCGCTACTGAAGGACAGCAGCTGCTGGGCTGCTGGGCTGACGCGGCAAGCTG  
ATCTGCACACCCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCTGGACAGATCTG  
GAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACCAACCTGATCTACA  
CCCTGATCGAGGAGGACGAGAACAGCAGGAGAAAGACGAGCAGGAGCTGCTGGAGCTGGA  
CAAGTGGGCCAGCCTGTGGAAGTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTT  
CATCATGATCGTGGGCGGCTGCTGGGCGCTGCGCATCGTGTTCACCGCTGCTGAGCATCGTGAA  
CCGCGTGCGCCAGGGCTACAGCCCTGAGCTTCCAGACCCGCTTCCCCGCCCCCGCGGGCCC  
CGACCGCCCCGAGGGCATCGAGGAGGAGGGCGCGGAGCGCGACCCGACCCGACGACGCCCC  
CTGCTGCAACGCGCTGCTGGCCCTGATGCTGGGACGACCTGCGCAGCCTGTGCTGTTTCACTAC  
CAGCGCTGCGCGACCTGATCTGATCGCGCCCGCATCTGTTGAGCTGCTGGGCGCGCGCGG  
TGGGAGGCCCTGAAAGTACTGGGGCAACCTGCTGCACTGAGTCTGGATCCAGGAGCTGAAGAACAG  
CGCGGTGAGCCTGTTGACGCGCATCGCATCGCGCTGGCGGAGGGCACCGACCGCATCATCGA  
GGTGGCCACGCGCATCGCGCGCCTTCTGCAATCCCCCGCGCATCCGCCAGGGCTTCGA  
GCGCGCCTGCTGTAACTCGAGCGTGCT

FIG. 10

## SEQ ID NO:8 VAL120-THR202

GAATTCCGCCACCATGGATGCAATGAAGAGAGGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA  
GTCTTCGTTTCGCCACGCGCGCTGGAGAAAGCTGTGGGTGACCGTGTAACACGCGGTGCCCGTG  
TGGAAAGAGGCCACACCAACCTGTTCTGCGCCAGCGACGCCAAGGCTTACGCACACCGAGGT  
GCACAACGTTGTGGGCCACCCACGCGCTGCGTGCCACCGACCCCAACCCCAAGGAGATCTGTGT  
GGAGAACGTCGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAGAGCCTGGAAGCCCTGCGTGGGGCGGCCACCCAGGCGCTG  
CCCCAAGGTGAGCTTCGAGCCCATCCCATCCACTACTGCGCCCCGCGCGCTTCGCCATCCT  
GAAATGCAACGCAAGAAGTTCAACGGCAGCGGCCCTGCAACCAACGTGAGCACCCTGCAAT  
GCACCAACGGCATCCGCCCCGTGGTGAGCACCAAGCTGCTGTAACGGCAGCCTGGCCGCTG  
GAGGGCGTGGTGATCCGACGCGAGAATTCACCGACAACGCCAAGACCATCATCTGTCAGCT  
GAAGGAGAGCTGGAGATCAACTGCACCCGCCCAACAACAACCCGCAAGAGCATCACCA  
TCGGCCCCGCGCGCGCTTCTACGCCACCGCGACATCATCGGCGACATCCGCCAGGCCCACT  
GCAACATCAGCGCGGAGAAATGGAACAACACCCCTGAAGCAGATCTGTGACCAAGCTGCAGGCC  
CAGTTTCGGCAACAAGACCATCGTGTTCAGCAGAGCAAGCGCGCGACCCCGAGATCGTGAT  
GCACAGCTTCAACTGCGCGCGGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCACACGAC  
CTGGAAACAACACCATCGGCCCAACAACAACCAACGGCACCATCAACCTGCCCTGCCGATCA  
AGCAGATCATCAACCGCTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCATCCGCGGC  
CAGATCCGCTGCAGCAACAATCAACGGCGCTGTGCTGACCCGCGACGGCGGCAAGGAGAT  
CAGCAACACCAACCGAGATCTTCGCGCCCGCGCGCGCGCATGCGCGCAACTGGCGCAGCG  
AGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCTGGGCGTGCGCCCCCAACCAAGGCCAAG  
CGCCGCGTGGTGACGCGGAGAAAGCGCGCGTGACCTGGGCGGCATGTTCCTGGGCTTCCTG  
GGCGCCCGCGGACACCATGGGCGCGCGAGCCTGACCTTGACCGTGACGGCCCGCCAGCT  
GCTGAGCGGCATGTGCAGCAGCAACAACCTGCTGGCGCCATCGAGGCCCGACAGCACCT  
TGCTGACGCTGACCCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGTGCTGGCCGTGGAGCGC  
TACCTGAAGAGCAGCAGCTGCTGGGCATCTGGGGTGCAGCGGCCAAGCTGATCTGCACCAAC  
CGCGTGGCTGGAAACGCCAGCTGGAGCAACAAGAGCCTGGACAGATCTGGAACAACATGA  
CCTGGATGGAGTGGGAGCGGAGATCGACAACCTACCAACCTGATCTACACCTGATCGAG  
GAGAGCCAGAACAGCAGGAGAGAAAGACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCA  
GCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCG  
TGGGCGCGCTGGTGGGCTGCGCATCGTGTTCACCGTGTGAGCATCGTGAAACCGCGTGGCGC  
AGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCGCGCGGCCGACCGCCCCG  
ACCGCATCGAGGAGGAGGGCGCGGAGCGCGACCGCGACCGCAGCAGCCCCCTGGTGACCG  
CCTGTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCTGTTCAGTACCAACCGCTGCG  
CGACCTGATCTGATCGCCCGCGCATCGTGGAGCTGCTGGGCCCGCGCGCTGGGAGGCCCT  
GAAAGTACTGGGGCAACCTGCTGCACTGATTCAGGAGCTGAAGAACAGCGCGCTGAGGCC  
TGTTGACAGCCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCCAGC  
GCATCGCGCGCGCTTCTGCACATCCCCGCCCATCCGCCAGGGCTTCGAGCGCGCCCTGC  
TGTAACTCGAG

FIG. 11

## SEQ ID NO:9 TRP427-GLY431

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA  
GTCTTCGTTTCGCCCAGCGCCGTGGAGAAAGCTGTGGGTGACCGGTACTACGGCGTGCCCGGTG  
TGGAAAGGGGCCACCAACCCCTGTTCTGCGCAGCGACGCCAAGGCCTACGCACAACCGAGGT  
GCACAACCTGTGTGGGCCACCCAGCCTGCTGTCGCCACCGACCCCAACCCCAAGGAGATCGTGTGCT  
GGAGAACGTGACCGGAGAATCTCAACATGTGGAAGAAACAACATGGTGGAGCAGATGACACGAG  
GACATCATCAAGCCTGTGGGACCAAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCTGCTGCGGTG  
ACCTTGCACTGGACCAACCTGAAGAAGCCCAACCAACCAAGAGCAGCAACTGGGAAGGAGAT  
GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAAGCATCCGCAACAAGATGC  
AGAAAGGAGTAGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAAGC  
TACAAGCTGATCAACTGCAACACCCAGCGTGATCACCCAGGCCTGCCCAAGGTGAGCTTGGAG  
GCCCATCCCCATCCACTACTGCGCCCGCCCGGCTTCGCCATCTCTGAAGTGCAACGACAAGAA  
GTTCAACCGGACGGGCCCTGCACCAACGTGAGCACCGTGCAAGTGCAACCCAGGCATCCCGCC  
CCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGCC  
AGCGAGAACTTCAACCGACAACGCCAAGAACCATCATCTGTGCAAGTGAAGGAGAGCGTGGAGAT  
CAACTGCACCCGCCCAACAACAACCCGCAAGAGCATCAACATCGGCCCGGCCCGCGCT  
TCTACGCCACCGCGACATCATCGCGGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG  
AAGTGGAAACAACCCCTGAAGCAGATCGTGACCAAGCTGCAGGCCAGTTTCGGCAACAAGAC  
CATCGTGTTCAAAGCAGAGCAGCGGCGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG  
GCGCGAGATTCTTCTACTGCAACAAGCACCAGCTGTTCAACAGCACCTGGAAACAACACCATCG  
GCCCAACAACACCAACGGCACCATCAACCTGCCCTGCCGCATCAAGCAGATCATCAACCGCT  
GGGGCGGCAAGGCCATGTAGCCCCCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACATC  
ACCGCGCTGCTGCTGACCCGCGAGCGCGGCAAGGAGATCAGCAACACCAACCGAGATCTTCGG  
CCCCGGCGCGCGACATCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGA  
AGATCGAGCCCTGGGCGTGGCCCCCAAGGCCAAGCGCCGCGTGGTGACGCGGAGAG  
CGCGCGGTGACCTGGGCGCATGTTCTGGGCTTCTGGGCGCGCGGCCAGCACCATGGGC  
GCCCAGACCTGACCTGACCGTGACCGTGACGGCCGCCAGCTGCTGAGCGGCATCGTGACGACGCA  
GAACAACCTGCTGCGGCCATCGAGGCCACGACGACCTGCTGCAGCTGACCGTGTGGGCA  
TCAAGCAGCTGCAGGCCCGCGTGTGCGCGTGGAGCGCTACCTGAAGGACAGCAGCTGCTG  
GGCATCTGGGGCTGCAGCGCAAGCTGATCTGCACACCGCCGTGCCGTGGAACGCCAGCTG  
GAGCAACAAGAGCCTGGACCAAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAG  
ATCGACAACCTACACCAACCTGATCTACAACCTGATCGAGGAGAGCCAGAACCAGCAGGAGAA  
GAAACGAGCAGGAGCTGCTGGAAGCTGGACAAGTGGGCGACCTGTGGAACCTGTTTCGACATCA  
GCAAGCTGTGTGTACATCAAGATCTTATCATGATCTGTGGGCGGCTGTTGGGCTGCGCA  
TGTGTTTACCGCTGTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCC  
AGAACCCTTCCCCGCCCGCGCGCCCGACCCGCCGAGGGCATCGAGGAGGAGGGCGCG  
GAGCGCAGCCGCAACCGCAGCAGCCCCCTGGTGACCGCCTGCTGGCCCTGATCTGGGACGA  
CCTGCGCAGCTGTGCTGTTCACTGCTACCAACCGCTGCGCGACCTGATCTGATCGCGCCCG  
CATCGTGGAGCTGCTGGGCGCGCGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGC  
AGTACTGGATCCAGGAGCTGAAGAAGCAGCGCTGAGGCTGTTGACGCCATCGCCATCGCC  
GTGGCCGAGGGCAACGACCGCATCATGAGGTGGCCAGCGCATCGGCGCGCGCTTCTGCA  
CATCCCCGCGCATCCGCCAGGGCTTCGAGCGCGCCTGCTGTAACTCGAG

FIG. 12

## SEQ ID NO:10 ARG426-GLY431

GAATTCCGCCACCATGGATGCAATGAAGAGAGGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA  
GTCTTCGTTTCGCCCCAGCGCGTGGAGAAAGCTGTGGGTGACCGGTGTACTACGGCGTGCCCGGTG  
TGGAAGGAGGGCCACCAACCCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGGAGT  
GCACAACGCTGTGGGCCACCCAGCGCTGCGTGCCCAACCGACCCCAACCCCGAGAGATCGTGTGCT  
GGAGAACGCTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCAAGAG  
GACATCAATCAGCCTGTGGGACCAAGCGCTGAAGCCCTGCGTGAAGCTGACCCCTCTGCGGTG  
ACCTTGCACTGCAACCAACTGAAGAACGCCACCAACCAAGAGCAGCAACTGGGAAGGAGAT  
GGACCGCGGGCGAGATCAAGAACTGACGCTTCAAGGTGACCAACGACATCCGCAACCAAGATGC  
AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGCTGGTGCCCATCGACAACGACACCAAGC  
TACAAGCTGATCAACTGCAACACCAAGCGTGATCAACCGAGGCTGCCCAAGGTGAGCTTCGA  
GCCCATCCCATTCACACTAGCGCCCGCGCGGCTTCGCCATCTGAAAGTGAACGACAAAGAA  
GTTCAACGGGACGCGGCCCTGCAACCAACGTGAGCACCGGTGCAAGTGAACCCACCGCATCCGCC  
CCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCGCTGGCCGAGGAGGGCGTGGTGATCCGC  
AGCGAGAACTTACCGGACAACGCCAAGACCATCATCTGTCAGCTGAAGGAGAGCGTGGAGAT  
CAACTGCAACCGGCCCAACAACAACCCGCAAGAGCATCACTCGGCCCGCGCGCGCT  
TCTACGCGACCGGGGACATCATCGGGGACATCGGCCAGGCCCATGCAACATACGCGGCGAG  
AAGTGAACAACAACCTGAAGCAGATCGTGACCAAGCTGCAAGGCCAGTTTCGGCAACAAGAC  
CATCGTGTTCGAAGCAGAGCAGCGCGGCGACCCCGAGATCGTGAAGCAGAGTTCACCAAGTGC  
GCGGCGAGTTCCTACTGCAACAACGACCCAGCTGTTTCAACAGCACCTGGAACAACCACTGCG  
GCCCAACAACAACCGGACCATCAACCTGCGCTGCCGATCAAGCAGATCATCAACCGC  
GGCGCGGCAAGGCCATGTATGACGCCGCCCATTCGCGGCCAGATCCGCTGACAGCAGCAACAT  
CACCGGCTGCTGCTGACCGCGGACGCGGCGCAAGGAGATCAGCAACAACCAACGAGATCTTCC  
GCCCCGCGCGCGCGGACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTG  
AAGATCGAGCCCTGGCGGTGGCCCCCACCAGGCCAAGCGCCGCTGGTGACGCGCGAGAA  
GCGCGCGGTGACCTGGGCGCCATGTTCTGGGCTTCTGGGCGCGCGCGGACGACCATGGG  
CGCCGCGAGCTGACCTGACCGGTGACGCGCGCGCGAGCTGCTGAGCGGCGATCTGTGACGAGC  
AGAACAACCTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGCACTGACCGTGTGGGCG  
ATCAAGCAGCTGCAAGGCCGCGTGTGCGCGGTGAGCGCTACCTGAAGGACGACGAGCTGCT  
GGGCATCTGGGGGTGCAAGCGGCAAGCTGATCTGCAACCGCGGTGCCCTGGAAACGCGAGCT  
GGAGCAACAAGAGCTGGACCAAGTCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGA  
GATCGACAACATACACCAACCTGATCTACACCTGATCGAGGAGAGCGCAAGCCAGCAGGAGA  
AGAAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCTGGGCGGCTGGTGGGCGCTGCGC  
ATCGTGTTCACCGTGTGAGCATCGTGAACCGCGTGGCGCAGGGCTACAGCCCCCTGAGCTTC  
CAGACCCGCTTCCCGCCCCCGCGGCCGACGCGCCGAGGGCATCGAGGAGGAGGGCGG  
CGAGCGCGACCGGACCGCGAGCGCGCCCTGCTGCAACGGCTGCTGGCCCTGATCTGGGACG  
ACCTGCGCAGCTGTGCTGTTCAGCTACACCGCTGCGCGAAGCTGATCTGATCGCGGCC  
GCATCGTGGAGCTGTGGCGCGCGCGCTGGGAGGGCCCTGAAGTACTGGGGCAACCTGCTG  
CGATGACTGATCCAGGAGCTGAAGAAACGCGCGCTGAGCCCTGTCGACGCATCGCCATCGC  
CTGGCGGAGGGGACCGGACCGCATCATCGAGGTGGCCAGCGCATCGGCGCGGCTTCTGCA  
ACATCCCCGCGCATCGCGCAGGGCTTGAGCGCGCCCTGCTGTAACCTGAG

FIG. 13



## SEQ ID NO:11 ARG426-GLY431B

GAATTGCGCCACCATTGGATGCAATGAAGAGAGGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA  
GTCTTCGTTTCGCCAGCGCGCTGGAGAGAAGCTGTGGGTGACCGGTGTACTACGGCGTGCCCGGTG  
TGGAAGGAGGCCACCAACCCCTGTTCTGGCCAGCGACGCCAAGGCTTACGACACCCGAGGT  
GCACAACGTGTGGGCCACCCAGCGCTGCGTGCCACCAGCCCAACCCCCAGGAGATCGTGTCT  
GGAGAACGTGTACCGGAGAACTTCAACATGTGGAAAGAAACAACATGTTGGAGCAGATGCACCGAG  
GACATCATCAGCCTGTGGGACCGAGAGCTGAAGGCCCTGCGTGAAGCTGACCCCCCTGTGCGTG  
ACCCTGCATGCAACCACTGAAGAAAGCCACCAACCAAGAGCAGCAACTGGAAGGAGAT  
GGACCGCGGCGGAGATCAAGAACTGCAGCTTCAAGGTGACCCAGCACTCCGCAACAAGATGC  
AGAAGGAGTACGCCCTGTTCTACAAGCTGGAGCTGTTGCCATCGCAACACGACAACACCCAGC  
TACAAGCTGATCAACTGCAACACCCAGCGTGATCACCAGGCCCTGCCCAAGGGTGAGCTTCGA  
GCCCATCCCACTCCACTACTGCGCCCCCGCGGCTTGCCCATCTGAAAGTGAACGACAAGAA  
GTTCAACGGCAGCGGCCCTGCAACAACGTGAGCAACCGTGCAAGTGCAACCCACGGCATCCGCC  
CCGTGGTGAGCACCAGCTGCTGTGAACGGCAGCCTGGCCGAGGAGGGCGTGTGTATCCGC  
AGCGAGAACTTCAACGACAACGCCAAGACCATCATCTGTCAGCTGAAGGAGAGCGTGAGAT  
CAACTGCACCCGCCCAACAACAACCCGCAAGAGCATCACCATCGGCCCGCGCGCGCT  
TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG  
AAGTGGAAACAACCCCTGAAGCAGATCTGTGACCAAGCTCGAGGCCAGTTCCGGCAACAAGAC  
CATCTGTGTTCAAGCAGAGCAGCGGCGGCGACCCGAGATCTGTATGCACAGCTTCAACTGCG  
GCGCGAGATGTTCTTACTGCAACAGCACCAGCTGTTCAACAGCACCTGGAAACAACACCATCG  
GCCCAACAACAACAGGCCACCATCACCCTGCCCTGCCCATCAAGCAGATCATCAACCCG  
GGCAGCGCAAGGCCATGTACGCCCCCCCACTCCGCGCCAGATCCGCTGCAAGCAACAT  
CACCGGCTGTCTGTGACCCGCGAGCGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCC  
GCCCGCGGCGGCGGACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTG  
AAGATCGAGCCCTTGGCGGTGGCCCCACCAAGGCCAAGCGCCGCTGTGTGACGCGCGAGAA  
GCGCGCGGTGACCTGGGCGCCATGTTCTTGGGCTTCTGGGCGCCGCCGACACCATGGG  
CGCCCGCAGCCTGACCTGACCGTGACGGCCCGCCAGCTGTGTAGCGGCCATGTGTGACGAGC  
AGAACAACCTGTGCGCGCCATCGAGGCCAGCAGCACCTGCTGAGCTGACCGGTGTGGGGC  
ATCAAGCAGCTGACGGCCCGCTGTGTGGCGTGAGGCGCTACCTGAAGGACCAGCAGCTGTCT  
GGGCATCTGGGGGTGACGCGCAAGCTGATCTGCACCAACCGCGTGCCTGGAACGCCAGCT  
GGAAGCAACAAGAGCTGGACAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGA  
GATCGACAACATACCAACCTGATCTACACCTGATCGAGGAGAGCCAGAACCAGCAGGAGA  
AGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACCTGGTTCGACATC  
AGCAAGTGGCTGTGGTACATCAAGATCTTCAATCATGATCGTGGCGGCCCTGGTGGGCTGCGC  
ATCGTGTTCACCGTGTGATGACATCGTGAACCGCGTGGCGCAGGGCTACAGCCCTGAGCTTC  
GAGACCCGCTTCCCGCCCCCGCGGCCCGACCCCGAGGGCATCGAGGAGGAGGGCGG  
CGAGCGCAGCCGACCGCAGCAGCCCGCTGGTGACCGGCTGTGTGGCCTGATCTGGGAGC  
ACCTGCGCAGCCTGTGCTGTTACGTACCAACCGCTGCGCGACCTGATCTGATCGCGGCC  
GCATCTGGAGCTGTGGGCGCGCGGCTGGGAGGCCCTGAAGTACTGGGCAACCTGTCTG  
CAGTACTGGATCCAGGAGCTGAAGAACAGCGCGCTGAGCGTGTTCGACGCCATCGCCATCGC  
CGTGGCCGAGGGCACCGACCGCATATCGAGGTGGGCCAGCGCATCGGCGCGCGCTTCTCTG  
ACATCCCCCGCCCATCCGCCAGGGCTTGAAGCGCGCCTGTGTAACTCGAG

FIG. 14

## SEQ ID NO:12 ARG426-LYS432

GAATTCCGCCACCA TGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA  
 GTCTTCGTTTTCGCCACGCGCGGTGGAGAAGCTGTGGGTGACCGGTGTACTACGGCGTGCCCGTGT  
 TGGAAAGGAGGCCACCAACCACTTCTTTCGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT  
 GCACAACCGTGTGGGCCACCAACGCTGCGTGGCCACCGACCCCAACCCCAAGGAGATGCTGTCT  
 GGAGAACCGTGAACCGAGAACTTCAACATGTGGAAGAACAAACATGGTGGAGCAGATGCACGAG  
 GACATCATCAGCCTGTGGGACAGAGCCTGAAAGCGCTGCGTGAAGCTGACCCCTGTGCGCTG  
 ACCCTGCACTGCAACCACTGAAGAAGCGCCACCAACACCAAGAGCAGCAATGGAAGGAGAT  
 GGACCGCGGGGAGATCAAGAACTGCAAGCTTCAAGGTGACCAACAGATCCCGCAACAAGATGC  
 AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCATCGACAAACGACCAACACCAAGC  
 TACAAGCTGTATCACTGCAACACAGCGTGTATCACCAAGCGCTGCCCAAGGTGAGCTTCGA  
 GCCCATCCCCATCCACTACTGCGCCCGCGCGGCTTCGCCATCCTGAAGTGCACCGACAAGAA  
 GTTCAACGGCGAGCGGCCCTGCAACCAACGTGAGCACCCTGCGAGTGCACCCACCGCATTCGCC  
 CGCTGGTGAACAACAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC  
 AGCGAGAACTTACCGGACAAACGCCAAGACCATCATCGTGCACTGAAGGAGAGCGTGGAGAT  
 CAACTGCACCCGCCCAACAACAACACCGCAAGAGCATCAACATCGGCCCGCGCCCGCCT  
 TCTACGCCACCGCGACATCATCGCGGACATCCGCCAGGCCCACTGCAACATCAGCGCGGAG  
 AAGTGGAAACAACCTGAAGCAGATCGTGACCAAGCTGCAGGCCAGTTCCGCAACAAGAC  
 CATCGTGTTCAGCAGAGCAGCGCGCGGAGACCCCGAGATCGTGATGCACAGCTTCAAAGTGC  
 GCGCGGAGTTCCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACTGGAAACAACACCATCG  
 GCCCAACAACAACCAACCGCACCATCAACCTGCCCTGCCCATCAAGCAGATCATCAACCGC  
 GCGCGGCAACAAGGCCATGTACGCCCGCCCATCCGCGGCCAGATCCGCTGCAGCAGCAACAT  
 CACCGCGCTGCTGTGACCCGCGACGGCGGCAAGGAGATCAGCAACACCAACCGAGATCTCC  
 GCCCGCGCGCGCGGACATGCGCGCAACAAGTGGCGACGCGAGCTGTACAAGTACAAGTGGTG  
 AAGATCGAGGCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCGCGTGGTGACGCGCGGAGAA  
 GCGCGCGGTGACCTGGGCGCATGTTCTTGGCTTCCTGGGCGCGCGCGGACGACCATGGG  
 CGCCCGCAGCGTGAACCTGACCGTGCAGGCCCGCCACGCTGCTGAGCGCATCTGTGACGAGC  
 AGAACAAACCTGTGCGCGCATCGAGGCCAGCAGCACTGCTGCAGCTGACCGTGTGGGGC  
 ATCAAGCAGCTGCAGGCCCGCGTGTGGCCGTGGAGCGCTAAGTGAAGGACAGCAGCTGCT  
 GGGCATCTGGGGTGCACGCGCAAGCTGATCTGCACCAACCGCGTGGCTGGAACGCCAGCT  
 GGAGCAACAAGAGCTGGACCAAGATCTGGAACAACATGAACCTGGATGGAAGTGGGAGCGCA  
 GATCGACAACACCAACCTGATCTACACCTGATCGAGGAGAGCGAGAACCAAGCAGGAGAG  
 AGAACGAGCAGGAGCTGCTGGAGCTGGACAAAGTGGGCCAGCTGTGGAACCTGTTTCGACATC  
 AGCAAGTGGCTGTGTACATCAAGATCTTATCATGATCGTGGGCGGCTGTTGGGCGTGGCG  
 ATCGTGTTCACCGTGTGAGCATCGTGAACCGCGTGCAGCGGCTACAGCCCCCTGAGCTTC  
 CAGACCCGCTTCCCGCCCCCGCGGCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGG  
 CGAGCGCGACCGGACCGCAGCAGCCCCCTGGTGCAACGGCTGCTGGCCTGATCTGGGACG  
 ACCTGCGCAGCCTGTGCTGTTTACGATACCAACCGCTGCGCGACCTGATCTGATCGCCGCC  
 GATCTGTGGAGCTGTGGGCGCGCGCGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTG  
 CAGTATCGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTTCGACGCCATCGCCATCGC  
 CGTGGCCGAGGACACGACCGCATCATGAGGTGGCCAGGCCATCGGCCGCGCTTCTCTGC  
 ACATCCCGCCGCGCATCCGCGAGGGCTTCGAGCGCGCCCTGCTGTAACCTGAG

FIG. 15

## SEQ ID NO:13 ASN425-LYS432

GAATTCGCCACCATGGATGCAATGAAGAGAGGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA  
GCTTCTCGTTTCGCCCCAGCGCCGTGGAGAAAGCTGTGGGTGACCGTGTACTACGGCGGTGCCGTGT  
TGGAAGGAGGGCCACACCAACCCCTGTTCTGCGCCAGCGGACGCCAAGGCCCTACGACACCGAGGT  
GCACAACCGTGTGGGGCACCCACGCGCTGCTGCCCCACCGACCCCAACCCCCAGGAGATCGTGCT  
GGAGAAACGTGACCGGAACTTCAACATGTGGAAGAAACAACATGTGTGAGCAGATGACGACGAG  
GACATCATCGACCTGTGGGACGAGAGCCTGGAAGCCTGCGTGAAGCTGACCCCTGTGCGGTG  
ACCCTGCACTGCAACAACTGAAGAACGCCACCAACCAAGAGCAGCAACTGGAAGGAGAT  
GGACCGCGGGCGAGATCAAGAAGTGCAGCTTCAAGGTGACCAACGACATCCGCAACAAGATGC  
AGAAGGAGTACGCCCTGTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACAGC  
TACAAGCTGATCAACTGCAACACCGAGGTGATCACCCAGGCCCTGCCCAAGGTGAGCTTCGA  
GCCCATCCCCATCCACTACTGCGCCCGCCGCGCTTCGCCATCCTGAAGTGCAACGACAAGAA  
GTTCAACGGGACGCGGCCCTGCACCAACGTGAGCACCCTGCGATGTCACCCACCGCATCCGCC  
CCGTGTTGAGCAACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCCGC  
AGCGAGAACTTCAACGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT  
CAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATACCATCGGCCCGGCCGCGCT  
TCTACGCCACCGCGACATCATCGGCGACATCCGCCAGGCCCATGCAACATCAGCGCGGAG  
AAGTGGAACAACACCTGAAGCAGATCGTGACCAAGCTGCAGGCCAGTTTCGGCAACAAGAC  
CATCTGTGTTCAAGCAGAGCAGCGCGCGGACCCCGAGATCGTGATGCACAGCTTCAACTGCG  
GCGCGGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACAACCATCG  
GCCCCAACAACCAACCGGACCATCAACCTGCCCTGCCGATCAAGCAGATCATCAACGCC  
CCAAGGCCATGTACGCCCGCCCCATCCGCGGCCAGATCCGCTGCAGCAGCAATCAACGGCC  
TGCTGTGACCCCGCAGCGCGGCAAGGAGATCAGCAACACCCAGAGATCTTCGCCCGCGG  
GGCGGCGACATGGCGGCAACAATGGCGCAGCGAGCTGACAAAGTGTGGTGAAGATCGA  
GCCCTGGGCGTGGCCCCCAACAGGCCAAGCGCGCGCTGTGTCAGCGCGAGAAGCGCGCG  
TGACCTGGGGCGCATGTTCTTGGGCTTCTGGGCGCGCGCGCAGCACCATGGGCGCGCGCA  
GCTTGACCTTGACCTGTCAGGCGCGGACGCTGCTGAGCGGCATCGTGACGACGAGAACAAC  
CTGCTGCGCGCCATCGAGGCCACGACGACCTGCTGTCAGCTGACCGGTGTGGGCGATCAAGCA  
GCTGCAAGGCCCGGTGCTGGCGCTGGAGCGCTACCTGAAGGACCAAGCAGCTGCTGGGCATCT  
GGGGTGCAGCGGCAAGCTGATCTGCAACACCGCGCTGCCCTGGAACGCCAGCTGGAGCAAC  
AAGAGCCTGGACCAAGATCTGGAAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAA  
CTACACCAACCTGATCTACACCTGATCGAGGAGAGGCCAAGCAACGACGAGAGAAGAACGAGC  
AGGAGCTGCTGGAGCTGGACAAGTGGGCCAGCCCTGTGGAACCTGGTTTCGACATACGCAAGTGG  
CTGTGGTATCATAGATCTTATCATATGATCGTGGGCGGCTGTTGGGCTGCGCATCTGTTTC  
ACCGTGCTGAGCATCTGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCCG  
TTCGCCGCCCGCGCGGCCGACGCCCGGAGGCACTGAGGAGGAGGGCGGCGAGCGGCA  
CCGCGACCGCAGCAGCCCCCTGTGTCACGGCCTGCTGGCCCTGATCTGGGACGACTGTGACG  
CCTGTGCCCTGTCACTACACCGCCTGCGCGAAGCTGATCCTGATCGCCGCCCGCATCTGTGGA  
GCTGCTGGGCGCGCGCGCTGGGAGGCCCTGAAAGTACTGGGGCAACCTGCTGCACTAGGA  
TCCAAGAGCTGAAGAACAGCGCGCTGAGCCTGTTTCAGCGCCATCGCCATCGCCGTGGCCGAG  
GGCACCCGACCGCATCATCGAGGTGGCCACGGCATCGGCCGCGCTTCTGCACATCCCCCG  
CGCATCCGCCAGGGCTTCGAGCGCGCCTGCTGTAACCTCGAG

FIG. 18

## SEQ ID NO:14 ILE424-ALA433

GAATTCGCCACCATGGATGCAATGAAGAGAGGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA  
GTCTTCGTTTCGCCCCAGCGCGCTGGGAGAAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG  
TGGAAAGGAGGGCCACCACCCCTGTTCTGCGGCCAGCGACGCCAAGGCCATCGACACCGCAGGT  
GCACAACGTGTGGGCCACCCACGCTGCGTGCCCAACGCCAACCCTCAGGAGATCGTGTCT  
GGAGAACGTGACCGGAGAATTCACATGTGGAAGAAACAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAAGAGCTGAAGCCCTGCGTGAAGCTGACCCCTCTGTGGTG  
ACCTTGACCTGACCAACCTGAAGAACGCCAACAACCAAGAGCAGCAATGGAAGGAGAT  
GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCAACAGCATCCGCAACAAGATGC  
AGAAAGGATACGCCCTGTCTACAAGCTGGACGTGGTGCCCATCGACAACGACAACACACG  
TACAAGCTGATCAACTGCAACACACAGCGTGATCACCCAGGCTGCCCAAGGTGAGCTTCGA  
GCCCATCCCCATCCACTACTGCGCCCCCGCGGCTTCGCCATCCTGAAAGTGCAACGACAAGAA  
GTTCACACGGCAGCGGCCCTGCACCAACGTGAGCACCGTGCAGTGACCCACGGCATCCCGCC  
CGTGCTGTGAGCACCCAGCTGCTGCTGAACGGCAGCTGGCCGAGGAGGGCGTGGTGATCCGC  
AGCGAGAATTCTACCGACAACGCCAAGACCATCATGCTGCAGCTGAAGGAGAGCGTGGAGAT  
CAACTGCACCCGCCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCCGCGCGCCT  
TCTACGCCACCGCGCATCATCCTCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG  
AAGTGGAACAAACACCTGAAAGCAGATCGTGACCAAGCTGCAGGCCAGTTTCGGCAACAAGAC  
CATCTGTTTCAAGCAGCAGCGCGCGGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG  
GCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCACAGCACCTGGAAACAACAACCATCG  
GCCCAACAACCAACCGGACCATCACCCTGCCCTGCCGATCAAGCAGATCATCGCGGCG  
GCACTGTATCGGCCGCCCATCCGCGGCGAGATCCGCTGCAGCAGCAACATACCCGCTGCTGT  
GTACCCCGCGACGCGGCAAGGAGATCAGCAACACCAACCGAGATCTTCGCGCCCGGCGCGGG  
CGACATGCGCGCAACTGCGCGCAGCGAGCTGTACAAGTACAAGTTGGTGAAGATCGAGCGCC  
TGGGCGTGCCCCCAAGGCCAAGCGCCGCTGGTGACAGCGCAGAGAAGCGCGCCGTGACC  
CTGGGCGCATGTTCTCTGGGCTTCTTGGGCGCCCGCGGACGACCATGGGCGCCCGCAGCCTG  
ACCTTGACCTGCGAGGCCCGCAAGCTGCTGAGCGGCATCGTGAGCAGCAGAAACAACCTGCT  
GCGCGCATCGAGGCCCAGCAGCACTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGC  
AGGCCCGGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAAGCAGCTGCTGGGCATCTGGGGC  
TGACGCGCAAGCTGATCTGCACCAACCGCGCTGCCCTGGAAACGCCAGCTGGAGCAACAAGAG  
CCTGGACCAAGATCTGGAACAACATGACCTGGATGGAAGTGGGAGCGCAGATCGACAACCTACA  
CCAACCTGATCTACACCTGATCGAGGAGAGGCCAGAACCAGCAGGAGAAGAACGAGCAGGA  
GCTGCTGGAGCTGGACAAGTGGGCCAGCCTGTGGAACTGGTTGACATCAGCAAGTGCGTGT  
GGTACATCAAGATCTTCAATCATGATCGTGGGCGGCGCTGGTGGGCTCGGCATCGTGTTCACCG  
TGCTGAGCATCGTGAACCGCGTGGCCAGGGCTACAGCCCTCGAGCTTCCAGACCCGCTTCC  
CGCGCCCGCGGCGCCGACCGCCCGGAGGCACTGAGGAGGAGGGCGGCGAGCGCGACCGC  
GACCCGACGACCCCTGGTGTGACGCGCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTG  
TGCTGTTCAGCTACCAACCGCTGCGCGACCTGATCTGATCGCGCCCGCATCTGTGAAGCTG  
CTGGGCGCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCACTGACTGATCA  
GGAGCTGAAGAACGCGCGGTGAGCTGTTGACGCCATGCCATCGCCGTGGCGAGGGCA  
CCGACCGCATCATCGAGGTGGCCAGCGCATCGGCCGCTTCTGCAATCCCCCGCGCA  
TCCGACAGGCTTCGAGCGCGCCCTGCTGTAACTCGAG

FIG. 17

## SEQ ID NO:15 ILE423-MET434

GAATTCGCCACCATGGATGCAATGAAGAGAGGGGCTCTGCTGTGTGCTGCTGTGTGTGGAGCA  
GTCTTCGGTTTCGCCACGCGCGTGGAGAAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG  
TGGAAAGGAGGCCACCACCAACCCCTGTTCTGCGCCAGCGACGCCAAGGCTACGACACCCGAGT  
GCACAAAGTGTGGGCGCACCCACGCGCTGGTGGCCACCGACCCCAACCCCGAGGAGATCTGTGCT  
GGAGAACGTGACCGGAGAATTCACATGTGGAAAGAACACATGTGTGGAGCAGATGCAACGAG  
GCATCATCAGCCTGTGGGACGAGAGCGCTGAAGCCCTGGGTGAAGCTGACCCCGCTGTGGGTG  
ACCTGCACCTGCACCAACCTGAAGAACGCCACCAACCAAGAGCAGCACTGGAAAGGAGAT  
GGACCGCGGCGAGATCAAGAACTGCAGCTTCAAGGTGACCACCAGCATCCGCAACAAGATGC  
AGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAAACGACCAACACCG  
TACAAGCTGATCAACTGCAACACCGCGTGATCACCCAGGCGCTGCCCAAGGTGAGCTTCGA  
GCCCATCCCATCCACTACTGGCGCCCGCGCGGCTTCGCCATCTGAAAGTGCAACGACAAAGAA  
GTTCAAACGGCAGCGGCCCTGCAACCAAGTGAGCACCGTGCAATGCAACCCACGGCATCCGCG  
CCGTGGTGAGCAACCCAGCTGCTGCTGAACGGCAGCGCTGGCCGAGGAGGCGGTGGTGATCCGC  
AGCGAAGACTTCACCGCAACCGCAAGACCATCATCTGCAGCTGAAGGAGAGCGTGGAGAT  
CAACTGCACCGCGCCCAACAACAACACCGCAAGAGCATCAACATCGGCCCGCGCGCGCT  
TCTACGCCACCGCGACATCATCGGCGACATCCGCGAGGCCACTGCAACATCAGCGCGGAG  
AAGTGGAAACAACACCTGAAGCAGATCGTGACCAAGCTGCAGGCCAGTTCGGCAACAAGAC  
CATGTGTTTCAACAGCAGAGCAGCGCGCGGCGACCCCGAGATCGTGATGCAACAGCTTCAACTGCG  
CGCGGCTTAAGTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAAACAACCATCG  
TGCCCAACAACAACAACCGGACCATCAACCTGCCCTGCCGATCAAGCAGATCGCGCGCATG  
TACGCCCGCCCATCCGCGCGAGATCCGCTGCAGCAGCAACATCACCGCGCTGCTGCTGACC  
CGCGACGGCGGCAAGGAGATCAGCAACACCACCGAGATCTTCGCCCGCGCGCGCGAGAT  
GCGCGACAACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCTGGGCG  
TGGCGCCACCAAGGCCAAGCGCGCGTGGTGAGCGCGAGAAAGCGCGCGTGAACCTGGGC  
GCCATGTTCTTGGGCTTCTTGGCGCGCGCGGCGAGCAACATGGCGCGCGCGAGCTGACCCCTG  
ACCGTGACGGCGCGCAGCTGCTGAGCGGATCGTGACGAGCAGAAACAACCTGCTGCGCG  
CATCGAGGCGCCAGCAGCACCTGCTGCAGCTGACCGGTGTGGGGCATCAAGCAGCTGCAGGCC  
GCGTGTGGCGGTGGAGCGCTACCTGAAGACCAAGCAGCTGCTGGGCACTGGGGCTGCAGC  
GGCAAGCTGATCTGCACACCGCGCTGCCCTGGAAACCGCAGCTGGAGCAACAAGAGCCTGGA  
CCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACTACACCAAC  
TGATCTACACCTGATCGAGGAGAGCCAGAACAGCAGGAGAGAAAGCAGCAGGAGCTGCTG  
GAGCTGGACAAGTGGGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCTGTGGTACAT  
CAAGATCTTATCATGATCGTGGCGGCGCTGGTGGCGCTGCGCATCGTGTTCACCGCTGTGAG  
CATGTGAAACCGCGTGTGCGCAGGGCTACAGCCCTGAGCTTCCAGACCGCTTCCCGCGCCC  
CGCGCGCCCGAGCCGCCCGAGGGCATCGAGGAGGAGGGCGCGAGCGCGACCCGCGACCGC  
AGCAGCCCGCTGTGTGACGCGCTGCTGGCGCTGATCTGGGACGACCTGGCGAGCGCTGTGCTG  
TTCAGCTACACCGCTGCGCGACCTGATCTGATCGCGCGCGCATCGTGGAGCTGCTGGCG  
CGCGCGCGTGGGAGGCGCTGAAGTACTGGGGCAACCTGCTGAGTACTGGATCCAGGAGCT  
GAAGAACACGCGCTGAGCGCTGTTCGAGCGCATCGCCATCGCGGTGGCGAGGGCAACGAGC  
GCATCATCGAGGTGGCCAGCGCATCGGCGCGCGCTTCTGCACATCCCCCGCGCATCCGCG  
AGGGCTTCGAGCGGCGCGCTGCTGTAACTCGAG

FIG. 18

## SEQ ID NO:16 GLN422-TYR435

GAATTCGCCACCATTGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGAGCA  
GTCTTCGTTTTCGCCACGCGCGTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGGCCGTG  
TGGAAAGGAGGCCACCACCACCTGTCTCTGCGCCAGCGACGCCAAAGGCTACGACACCCGAGGT  
GCACAACTGTGTGGGCCACCCACGCTGCGTGCCACCCGACCCCAACCCCAAGGAGATGTGTGCT  
GGAGAAGCTGACCGGAGAATCTCAACATGTGGAAAGAACACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAGAGCTGAAGCCCTGCGTGAAGCTGACCCCTGTGCGGTG  
ACCCTGCACTGCACCAACCTGAAGAAGCGCCACCAACACCAAGAGCAGCAACTGGAAGGAGAT  
GGACCGCGCGGAGATCAAGAATGCAAGCTTCAAGGTGACACCGATCCGCAACCAAGATGC  
AGAAAGGAGTACGCCCTGTCTACAAAGCTGGACGTGGTGCCCATCGACAACGACAACACCAAGC  
TACAAGCTGATCAACTGCAACACCGAGCTGATCACCCAGGCGTGCCTCAAGGTGAGCTTCGA  
GCCCATCCCCATCCACTACTGCGCCCCCGCGGCTTCGCCATCCTGAAAGTCAACAGCAAGAA  
GTTCACCGGCAGCGCCCCGCAACCAAGCTGAGCACCGTGCAGTGCACCCAGCGCATCCGCC  
CCGTTGGTGAGCACCCAGCTGCTGTCTGAAACGCGACGCTGGCCGAGGAGGGCGTGGTGATCCGC  
AGCGAGAACTTCAACGCAACACGCCAAGACCATCATCTGTCGAGCTGAAGGAGAGCGCTGGAGAT  
CAACTGCAACCCGCCCAACAACAACACCCGCAAGAGCATACCATCGGCCCCGCGCGCCCT  
TCTACGCCACCGCGGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGCGAG  
AAGTGGAAACAACACCTGAAGCAGATCGTGACCAAGCTGCAGGCCAGGTTCGGCAACAAGAC  
CATCTGTGTCAAGCAGACGACGCGGCGCGACCCCGAGATCGTGATGCACAGCTTCAACTGCG  
CGCGGTGGTTCTTCTACTGCAACAGCACCCAGCTGTTCACAGCACCTGGAAACAACACCATCG  
GCCCAACAACACCAACCGGCACCATCACCCTGCCCTGCCGTCATCAAGCAGGCGGCTACGCC  
CCCCCATCCGCGCGCAGATCCGCTGCAGCAGCAACATCACCGGCTGTGCTGTGACCCGCGAC  
GGCGGCAAGGAGATCAGCAACACCAACCGAGATCTTCCGCCCGCGCGCGCGACATGCGCGA  
CAACTGGCGCAGCGAGCTGTACAAGTACAAGTGGTGAAGATCGAGCCCTGGGCGTGGCCC  
CCACCAAGGCCAAGCGCCGCTGGTGACGCGGAGAAAGCGCGCCGTGACCTGGGCGCCATG  
TTCCTGGGCTTCTTGGCGCGCGCGCGCAGCACCATGGGCGGCCGACGCTGACCTGACCGGTG  
CAGGCCCGCAGCTGTGAGCGGCATCGTGACGACGAGAAACAACCTGCTGCGCGCCATCGA  
GGCCACGACGACCTGCTGCAAGTGCAGCTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGC  
TGGCCGTGGAGCGCTACCTGAAGGACGACAGCTGCTGGGCACTTGGGGCTGCAGCGGCAAG  
CTGATCTGCACCAACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGAGCTGGACCAAGAT  
CTGGAAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACAACCTACACCAACCTGATCT  
ACACCTGATCGAGGAGAGGCCAGAACGACGAGGAGAAAGACGACGAGGAGCTGCTGGAGCT  
GGAACAAGTGGCCAGCCTGTGGAACTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGA  
TCTTCATCATGATCTGGGCGGCTGTGGGCTGCGCATCGTGTTCAACGCTGCTGAGCATCG  
TGAACCGCGTGCAGCGCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCGCCCCCGCG  
GCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGCGCAGCGCGACCCGACCGCAGCAGAC  
CCCCCTGGTGACGCGCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCTGTTTCAAG  
CTACCAACCGCTGCGCGACCTGATCTGATCGCCGCCCGCATCTGGAGAGCTGTGGGCGCGCG  
CGGCTGGGAGGCCCCGAAGTACTGGGGCAACCTGCTGCACTGAGTATCGGATCCAGAGCTGAAGA  
ACAGCGCGTGAGCCTGTTGACGCGCATCGCCATCGCGTGCGGCGAGGCGACCCGACCGCATC  
ATCGAGGTGGCCACGCGATCGGCCCGGCTTCTTGCACATCCCCCGCGCATCCGCCAGGGC  
TTCGAGCGCGCCCTGTGTAACTCGAG

FIG. 19

## SEQ ID NO:17 GLN422-TYR435B

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA  
GTCTTCGTTTCGCCCCAGCGCGCTGGAGGAAGCTGTGGGTGACCGTGTACTACGGCGTGGCCGTG  
TGGAAAGGAGGCCACCAACCCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT  
GCACAACTGTGTGGGCCACCCAGCTGCGTGCCCAACGACCCCAACCCCAAGGAGATCGTGTGT  
GGAGAACCGTGAACCGGAACCTTCAACATGTGGAAGAAACAACATGGTGGAGCAGATGCAACGAG  
GACATCATAGCCTGTGGGACGAGAGCTGAAGCCCTGCGTGAAGCTGACCCCTGTGCGGTG  
ACCTTGCACTGCAACCACTGAAGAAGCGCAACCAACCAAGAGCAGCAACTGGAAAGGAGAT  
GGACCGCGGGCAGATCAAGAAGCTGCAGCTTCAAGGTGACCAACAGCATCCGCAACAAGATGC  
AGAAAGGATACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATGCACAACGACAACACCGAGC  
TACAAGCTGATCAACTGCAACACCAAGCGTGATCAACCAAGCGCTGCCCAAGGTGAGCTTCGA  
GCCATCCCCATCCACTACTGCGCCCGCGCGGCTTCGCCATCTGAAAGTGCAACGACAAAGAA  
GTTCAACGGCAGCGGCCCTGCACCAACGTGAGCACCGTGCAGTGCACCCACGGCATCCCGC  
CCGTGGTGACACCAAGCTGCTGCTGAACGGCAGCGCTGGCGGAGGAGGGCGTGGTGATCCCG  
AGCGAGAACTTACCGGACAACGCCAAGACCATCATCGTGCAGCTGAAGGAGAGCGTGGAGAT  
CAACTGCACCCGCCCAACAACAACCCCGCAAGAGCATCAACATCGGCCCGCGCGCGCT  
TCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCAGCGGGGAG  
AAGTGAAACAACCCCTGAAGCAGATCGTGACCAAGCTGCAAGGCCAGTTCGGCAACAAGAC  
CATCGTGTTCAGCGAGCAGCGCGCGGACCCCGAGATCGTGATGCACAGCTTCAACTGCG  
GCGGCGAGTCTTCTACTGCAACAGCACCAGCTGTTTAAACAGCACCTGGAACAACACCATCG  
GCCCAACGAGAACCAACCGCACCATCAACCTGCCCTGCGCATCAAGCAGGCCCTCATCGCC  
CCCCATCCCGCGGCAGATCCGCTGCAGCAGCAACATCAACCGGCTGTGTGTGACCCGCGAGC  
GGCGCAAGGAGATCAGCAACCAACCGAGATCTTCGCCCGCGCGCGCGGACATGCGCGAC  
AACTGGCGCAGCGAGCTGTACAAGTACAAGTGGTGGTGAAGATCGAGCCCTGGCGCTGGCCCC  
CACCAAGGCCAAGCGCGCGTGGTGACGCGGAGAAGCGCGCTGACCTGGCGGCATGT  
TCCTGGGCTTCTGGGCGCGCGCGGAGCACCATGGCGCGCGCGCTGACCTGACCTGACCGTGC  
AGGCCCGCGAGCTGTGAGCGGACATCGTGACGAGCAGACAACCTGCTGCGCGGCATCGAG  
TGATCTGCACCAACCGCGTGCCTGGAAACGCCAGCTGGAGCAACAAGACCTGGACCAAGATC  
TGGAAACAATGACCTGGATGGATGGGAGCGCGAGATCGACAACATCAACAACCTGATCTCA  
CACCTGATCGAGGAGAGCCAGAAACAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTG  
GACAAAGTGGGCCAGCTGTGGAACGTGTTTCAGATCAGCAAGTGGCTGTGTATCAATCAAGAT  
CTTCACTATGATCGTGGGCGGCTGGTGGGCTGCGCATCGTGTTCACCTGTGAGCATCGT  
GAACCGCGTGGCGCAGGGCTACAGCCCTTGAGCTTCAGAACCCGCTTCCCGCGCGCGCGCGG  
CCCCGACGCGCCAGGGCATCGAGGAGGAGGCGCGGAGCGCGACCGCGACCGCAGCAGCAGC  
CCCCGTGTGTGACGGCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCGCTGTGCTGTTCAGC  
TACACCGCTGCGCGAAGCTGATCTGATCGCCCGCGCATCGTGGAGCTGTGGGCGCGCGC  
GGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGCTGCACTGATGATCCAGGAGCTGAAGAA  
CAGCGCGCTGAGCTGTTCGACGCCATCGCCATCGCGTGGCGGAGGCGACCGACCGCATCAT  
CAGAGTGGCCACCGCATCGGCCGCGCTTCTGCATCCCGCGCGCATCGCGAGGGCT  
CGAGCGCGCTGCTGTAACTCGAG

FIG. 20

## SEQ ID NO:18: LEU122-SER199; ARG426-GLY431

GAATTCCGCCACCATGGATGCAATGAAGAGAGGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA  
GTCTTCGTTCGCCACGCGCTGGGAGAAGCTGTGGGTGACCGTGTAACACGCGTGCCCGTG  
TGGAAAGGAGGCCACCAACACCTGTCTGCGCCAGCGACGCCAAGGCCTACGCAACCGAGGT  
GCACAACCGTGTGGGCCACCAACGCTGCGTGCCCAACGACCCCAACCCCAAGGAGATCGTGTG  
GGAGAACGTGACCGGAGAATTCACATGTGGAAGAAACAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCTGTGGGACGAGCCTGAAGCCCTGCGTGAAGCTGGGGCAACAGCGTGAT  
CACCCAGGCCTGCCCCAAGGTGAGCTTCGAGCCCATCCCATCCACTACTGCGCCCCCGCGG  
CTTCGCCATCTGGAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCACCAACGTGA  
GCACCGTGCACTGCACCCACGGCATCCGCCCGCTGGTGAGCACCCAGCTGCTGCTGAACGGC  
AGCCTGGCCGAGGAGGGCGTGGTGATCCGCAAGGAACTTCACCGACAACGCCAAGACCAT  
CATCTGCAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCA  
AGAGCATACCATCGGCCCGGCCGCGCCTTCTACGCCACCGCGACATCATCGGCACATCC  
GCCAGGCCACTGCAACATCAGCGCGGAGAAGTGGAAACAACACCTGAAGCAGATCTGTGACC  
AAGCTGCAGGCCAGTTCGGCAACAAGACCATCGTGTTCAAGCAGAGCAGCGCGCGGCAACC  
CGAGATCTGTATGCACAGCTTCAACTGCGCGCGCGAGTTCTTCTACTGCAACAGCACCCAGCT  
GTTCAACAGCACTGGAAACAACACCATCGGCCCAACAACAACCAACGGCACCATACCCCTGC  
CCTGCCGATCAAGCAGATCATCAACCGCGCGCGCGCAAGGCCATGTACGCCCCCCCCATCC  
GCGGCCAGATCCGCTGCAGCAGCAACATCACCGGCTGCTGCTGACCCGCGCAGCGCGGCAAG  
GAGATGCAGCAACACCAACCGAGATCTTCGCCCGCGCGCGCGGACATGCGCGACAACATGGCG  
CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCAGGCCCTGGGCGTGCCGCCCAACCAAGG  
CCAAGCGCGCGGTGGTGACGCGGAGAAGCGCGCGGTGACCTGGGCGCCATGTTCTCGGGC  
TTCCTGGCGCGCGCGGACGACCATGGGCGCCGCGAGCCTGACCTGACCGTGCAGGCCCGC  
CAGCTGCTGAGCGGCATCTGTCAGCAGCAGAACAACCTGCTGCGGCCATCGAGGCCAGCA  
GCACCTGCTGACGTGACCGGTGTGGGCAATCAAGCAGCTGCAGGCCCGGTGCTGGCGGTGG  
AGCGCTACTGAAGGACCAAGCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTCG  
ACCACCGCCGTGCCCTGGAACGCCAGCTGGAGCAACAAGACCTGGACCAAGATCTGGAACAA  
CATGACCTGGATGGAGTGGGAGCGGAGATCGACAACCTACCAACCTGATCTACACCTGCA  
TCGAGGAGAGCCAGAACAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAAGCTGGACAAGTG  
GGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAAGTGGCTGTGGTACATCAAGATCTTCATCAT  
GATCGTGGGCGGCTGGTGGGCTGCGCATCGTGTTACCGGTGCTGAGCATCGTGAAACCGGT  
GCGCCAGGGCTACAGCCCTTGAGCTTCCAGACCCGCTTCCCCGCCCGCGCGGCCCGGACCG  
CCCCGAGGGCATCGAGGAGGAGGGCGGCGAGCGCGACCGCGACCGCAGCAGCCCTGTGTGC  
ACGGCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCTGTTCAGTACCAACCGCC  
TGGCGCAGCTGATCTGATCGCCGCCCGCATCTGTGAGCTGCTGGGCGCGCGGGCTGGGAGG  
CCCTGAAGTACTGGGGCAACCTGCTGCACTGAGTCTGAGTCCAGGAGCTGAAGAACAGCGCCGTG  
AGCCTGTTTCAGGCCATCGCCATCGCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC  
CAGCGCATCGCGCGCCCTTCCTGCACATCCCCCGCGCATCCGCCAGGGCTTCGAGCGCGCC  
CTGCTGTAACTCGAG

FIG. 21



## SEQ ID NO:19 LEU122-SER199; ARG426-LYS432

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA  
GTCCTCGTTTCGCCCCAGCGCCGTGGAGAAAGCTGTGGGTGACCGGTACTACGGCGTGCCCGGTG  
TGGAAAGGAGGGCCACCACCACCTGTTCTGCGCCAGCGACGCCAAGGCCATACGACACCGAGGT  
GCACAACGTGTGGGCCAACCCAGCTGCGTGGCCACCGACCCCAACCCCAAGGAGATCGTGCT  
GGAGAACGTGACCGGAACTTCAACATGTGGAAAGAACAAATGTTGGAGCAGATGCACGGAG  
GACATCATCAGCCTGTGGGACCCAGAGCTTGAAGCCTCGTGGAAGCTGGGCAACAGCTGTGAT  
CAGCCAGGCCGTGCCCAAGGTGAGCTTGGAGCCATCCCATCCACTACTGCGCCCCCGCGG  
CTTCGCCATCCTGAAAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCAACCAACGTGA  
GCACCGTGCAGTGCACCCACGGCATCCGCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGC  
AGCCTGGCCGAGGAGGGCGTGGTGATCCGACGCGAGAACTTCAACGACAACGCCAAGACCAT  
CATCTGTCAGCTGAAGGAGAGGTGGAGATCAAACCTGCACCGCCCCAACAAACAACCCGCA  
AGAGCATACCATCGGCCCGGCCGCGCTTCTACGCCACCGCGACATCATCGCGACATCC  
GCCAGGCCCACTGCAACATCAGCGGCGAGAAAGTGGAACAACAACCTGAAGCAGATCGTGACC  
AAGCTGCAGGCCCAAGTTCCGGCAACAAGACCATCGTGTTCAGCAGAGCAGCGCGCGGACCC  
CGAGATCTGTGATGCACAGCTTCAACTGCGCGCGCGAGTTCTTCTACTGCAACAGCACCCAGCT  
GTTCAACAGCACTGGAAACAACCATCGGCCCAACAACAACCAACGCCACCATCACCTGCG  
CCTGCGGCATCAAGCAGATCATCAACCGCGCGCGCAACAAGGCCATGTACGCCCCCCCCATCC  
CGCGCCAGATCCCTGCAGCAGCAACATCAACGGCCTGCTGTGACCCGCGCAGCGCGGCAAG  
GAGATGACGAACAACCCAGAGATCTTCGCCCGCGCGCGCGAGACATGCGCGCAACTGGCG  
CAGCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCTGGGCGTGGCCCCCAACCAAG  
CCAAGCGCCCGTGTGTCAGCGCGAGAAGCGCGCGCTGACCTGGCGCCATGTTCCTGGGC  
TTCCTGGGCGCGCCGCGCAGCACCATGGCGCGCCGACGCTGACCTGACCGTGACGGCCCGC  
CAGCTGCTGAGCGGCATCGTGACGACGAGAAACAACCTGCTGCGCGCCATCGAGGCCACGCA  
GCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGGTGCTGGCCGTGG  
AGCGCTACCTGAAGGACAGCAGCTGCTGGGCATCTGGGCTGCAGCGGCAAGCTGATCTGCG  
ACCACCGCCGTGCCCTGGAAAGCCAGCTGGAGCAACAAGAGCCTGGACAGATCTGGAACAA  
CATGACCTGGATGGAGTGGGAGCGCGAGATCGCAACATACACCAACCTGATCTACACCTGTA  
TCGAGGAGAGCCAGAACCAAGCAGGAGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTG  
GGCCAGCCTGTGGAACCTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCAT  
GATCGTGGGCGCGCTGGTGGGCTGCGCATCGTGTTCACCGTGCTGAGCATCGTGAACCGCGT  
GGCCAGGGCTACAGCCCTGAGCTTCCAGACCCGCTTCCCGCCCCCGCGGCCCGCAGCG  
CCCCAGGGCATCGAGGAGGAGGGCGCGAGCGCGACCCGACCCGACAGCCCCCTGGTGC  
ACGGCCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCTGTTCAGCTACCAACCGCC  
TGCGCGACCTGATCCTGATCGCCGCGCGCATCTGTGGAGCTGTGGGCGCGCGCGCTGGGAGG  
CCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTG  
AGCCTGTTTGAAGCCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGCC  
CAGCGCATCGGCCGCGCCTTCTCGCATCCTCCCGCGCGCATCGCCAGGGCTTCGAGCGCGCC  
CTGCTGTAACTCGAG

FIG. 22

## SEQ ID NO: 20: LEU122-SER199; TRP427-GLY431

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGTGGAGCA  
GTCTTCGTTTCGCCACGCGCTGGAGAAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG  
TGGAAAGGAGGCGCACCAACCCCTGTTCTGCGCCAGCGAGCGCAAGGCTACGACACCGAGGT  
GCACAAACGTGTGGGCGACCAACGCTGCGTGCACCGCAACCCCAAGGAGATCGTGCT  
GGAGAAAGTGTACCGGAGAACTTCAACATGTGGAGAGAAACATGGTGGAGCAGATGCACGAG  
GACATCATTCAGCCTGTGGGACCAAGAGCTGAAGCCCTGCGTGAAGCTGGGCAACAGCGTGAT  
CACCACAGGCTGCCCCAAGGTGAGCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCCGG  
CTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAAOGGACGCGCCCCCTGCACCAACGCTGA  
GCACCGTGCAGTGCACCCACGCGATTCGCCCCGTGGTGAGCACCCAGCTGCTGTGTAACGCGC  
AGCCTGGCGGAGGAGGGCGGTGTGTATCCGAGCGAGAACTTACCGACAAACGCCAAGACCAT  
CATCTGTGAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCA  
AGAGCATACCATCGGCCCGGCCGCGCTTCTACGCCACCGGCGACATCATCTGGCGACATCC  
GCCAGGCCCACTGCAACATCAGCGGCGGAGAAGTGGAAACAACCCCTGAAGCAGATCGTGACC  
AAGCTGCAGGCCAGTTTGGGCAACAAGACCATCGTGTTCAGCAGAGCAGCGCGCGCGACCC  
CGAGATCGTGATGCACAGCTTCAACTGCGGCGGCGAGTCTTCTACTGCAACAGCACCCAGCT  
GTTCACAAGCACTGGAAACAACACCATCGGCCCAACAACACCAACGGACCATCACCTGCG  
CTGCGCGCATCAAGCAGATCATCAACGCTGGGGCGGCAAGGCCATGTACGCCCCCCCCATCC  
CGCGGCAGATCCGCTGCAGCAGCAACATCACCGGCTGCTGCTGACCCGCGACGCGCGGCAAG  
GAGATCAGCAACACCAACCGAGATCTTCCGCCCCGCGCGCGGACATGCGCGCAACAATGGGG  
CAGCGAGCTGTACAAGTACAAGTGGTGAAGATCGAGGCCCTGGGCGTGGCCCCCAACGAAG  
CCAAGCGCGCGGTGTGTCAGCGCGGAGAAGCGCGCGTGAACCTGGGCGCCATGTTCTTGGCG  
TTCCTGGCGCGCGCGGACGACCATGGCGCGCCGAGCCTGACCCCTGACCGTGCAGGCCCGC  
CAGCTGCTGAGCGGCATCGTGAGCAGCAGAAACCTGCTGGCGCGCATCGAGGCCGAGCA  
GCACCTGCTGCAGCTGACCGTGTGGGCAATCAAGCAGCTGCAGGCCCGCGTGTGGCGCGTG  
AGCGCTACCTGAAGGACGAGCAGCTGCTGGGCACTTGGGGCTGCAGCGGCAAGCTGATCTGC  
ACCACCGCGTGGCCCTGGAAACGCCAGCTGGAGCAACAAGAGCCTGGACCAAGATCTGGAACAA  
CATGACCTGGATGGAGTGGGAGCGGAGATCGACAACATCAACCAACCTGATCTACACCTGTA  
TCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTG  
GGCCAGCCTGTGAACTGGTTCGACATCAGCAAGTGGCTGTGTTACATCAAGATCTTATCAT  
GATCGTGGCGCGCTGGTGGGCTGCGCATCTGTTCACCGTGTGAGCATCGTGAACCGCGT  
GCGCCAGGGCTACAGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCGCGCGGCCGACCG  
CCCCGAGGACATCAGGAGGAGGCGCGGAGCGGACCGGACCGCAGCAGCCCCCTGGTGC  
ACGGCTGCTGGCCCTGATCTGGGACGACCTGCGCAGCCTGTGCTGTTCAGTACCAACCGCC  
TGCAGCAGCTGATCTGATCGCGCGCGCATGTGTGAGCTGCTGGGCGCGCGCGGCTGGGAGG  
CCCTGAAGTACTGGGGCAACCTGCTGCGAGTCTGGATCCAGGAGCTGAAGAACACGCGCGGT  
AGCCTGTTTCGACGCCATCGCCATCGCGCTGGCCGAGGGCACCGGACCGCATCATCGAGGTGGC  
CAGCGCATCGCGCGCCTTCTGACATCCCCGCGCATCCGCCAGGGCTTCGAGCGCGCC  
CTGCTGTAACTCGAG

FIG. 23

SEQ ID NO:21 LYS121-VAL200; ASN425-LYS432

GAATTGCCACCATTGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA  
 GTCTTCGTTTCGCCACGCGCGTGGAGAAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG  
 TGGAAAGGAGGCGACACCAACCCCTGTTCTGCGCCACGCGACGCCAAGGCCATACGACAACCGAGGT  
 GCACAACGCTGTGGGCCACCCACGCTGCGTGGCCACCGACCCCAACCCCAAGGAGATCGTGCT  
 GGAGAACGTGACCGGAGAATTCAACATGTGGAAAGAACAAATGTGTGGAGCAGATGTCACGAG  
 GACATCATCAGCCTGTGGGACCAAGAGCCTGGAAGCCCTGCGTGAAGGCCCGCTGTATCACCCA  
 GGCCTGCCCAAGGTGAGCTTCGAGGCCATCCCATCCACTACTGCGCCCCCGCCGGCTTCGC  
 CATCTCGAAGTGCACGCAAGAAAGTTCAACGGCAGCGGCCCTGCAACAACGTGAGCAACCG  
 TGCAGTGACCCACGGCATCCGCCCGCTGGTGAGCACCCAGCTGCTGCTGAACGCGAGCCTGG  
 CGAGGAGGGCGTGGTGATCCGACAGGAGAACTTCAACGCAACGCCAAGAACCATCATCTGTG  
 CAGCTGAAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCAT  
 CACCATCGGCCCGCGCCGCGCTTCTACGCCACCGCGCAGCATCATCGCGCATATCCGCCAGGC  
 CCACTGCAACATCAGCGGCGAGAAAGTGGAACAACACCCCTGAAGCAGATCTGTGACCAAGCTGC  
 AGGCCAGTTCGGCAACAAGACCATCTGTGTTCAAGCAGAGCAGCGGCGCGACCCCGAGATC  
 GTGATGCAACAGCTTCAACTGCGGCGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTTCAAC  
 AGCACTGGAAACAACACCATCGGCCCAACAACAACAACGCGACCATCAACCTGCCCTGCCG  
 CATCAACGAGATCATCAACGCCCAAGGCCATGTACGCCCGCCCATCCGCGGCGAGATCCG  
 CTGACGAGCAACATCAACCGCCTGCTGCTGACCCGCGACGCGCGCAAGGAGATCAGCAACA  
 CCACCGAGATCTTCGCCCGCGCGCGCGGACATGCGCGACAACCTGGCGAGCGAGCTGTAC  
 AAGTACAAGGTGTGTGAAGATCGAGCCCTGGGCGTGGCCCCCAACAAGGCCAAGCGCCGCT  
 GTGTGACGCGGAGAAAGCGCGCGTGAACCTGGGCGCATGTTCTGGGCTTCTGGGCGCGCG  
 CGGCGACCATGGGCGCGCGGAGCCTGACCCCTGACCGGTGACGGCCCCGCGAGCTGTGAGCG  
 GCATCGTGACGACGAGAAACAACCTGTGCGCGCATCGAGGCCACGACGACTGCTGTGAG  
 CTGACCGTGTGGGGCATCAACAGCTGACAGGCCCGCGTGTGCGCGTGGAGCGCTACTGTAA  
 GGACAGCAGCTGTGGGCATCTGGGCTGACGCGGCAAGCTGATCTGCACACCGCCGTGC  
 CTTGGAACCGCACTGGAGCAACAAGAGCCTGGACCATCTGGAAACAACATGACCTGGAATG  
 GAGTGGGAGCGCGAGATCGACAACTACCAACCTGATCTACACCTGATCGAGGAGAGCCA  
 GAACCAAGCAGGAGAAAGACGAGCAGGAGCTGTGAGCTGGACAAAGTGGGCGACCTGTGG  
 AACTGGTTCAGCATCAGCAAGTGGCTGTGGTACATCAAGATCTTCAATCATGATCGTGGGCGGC  
 CTGTGTGGGCTGCGCATCGTGTTCACCGTGTGTGAGCATCGTGAACCGCGTGGCCAGGGCTAC  
 AGCCCCCTGAGCTTCCAGACCCGCTTCCCCGCCCGCGCGGCCGACCGCCCCGAGGGCATC  
 GAGGAGGAGGGCGCGGAGCGCGAACCGGACCGCAGCAGACGCCCTGTGTGACGCGCTGTGGC  
 CCTGATCTGGGACGACCTGCGCAGCCTGTGCTGTTCACTACCAACCGCCTGGCGACCTGAT  
 CTTGATGCGCGCCCATCTGTGAGCTGTGTGGCGCGCGCGCTGGGAGGCCCTGAAGTACTG  
 GGGAACCTGCTGAGTACTGGATCCAGGAGCTGAAAGAACAGCGCCGTGAGCCTGTTCAGCG  
 CCATCGCCATCGCGGTGGCGGAGGGCACGACCGCATCATCGAGGTGGCCCAAGCGCATCGGC  
 CGCGCTTCTGCACATCCCCGCCCATCCGCCAGGGCTTCGAGCGCGCGCTGTGTAACTC  
 GAG

FIG. 24

SEQ ID NO:22 VAL120-ILE201; ILE 424-ALA433

GAATTCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA  
GTCCTTCGTTTGGCCACGCGCTGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG  
TGGAAAGGAGGGCCACCACCCCTGTTCTGCGCCAGCGACGCCAAAGGCCTACGCACACCGAGGT  
GCAACAACGCTGTGGGCCACCCACGCTGCGTGCCACCGAACCCCAACCCCAAGGAGATCGTGCT  
GGAGAACGTGACCGAGAACTTCAACATGTGGAAGAAACACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCGAGGCTGAAGCCCTGCGTGGGCGGCATCACCCAAGCCTG  
CCCCAAGTGTAGCTTCGAGCCCATCCCATCCACTCTGCGGCCCGCGCGCTTCGCCATCCT  
GAAGTGCAACGACAAAGAAGTTCAACGGCAGCGGCCCTGCACCAACGTGAGGCAACCGTGACGT  
GCACCCACGGCATCCGCCCGTGTGTGAGCACCAGCTGCTGCTGAACGGCAGCCTGGCCGAG  
GAGGGCGTGTGATCCGCAGCGAGAACTTCAACGGCAACGCCAAGACCATCATCTGTGACGT  
GAAGGAGAGCGGTGGAGATCAACTGCAACCGCCCCAACAAACACCCGAAAGAGCATCACCA  
TCGGCCCCGGCCGCGCTTCTACGCCACCGCGGCATCATCGGCGACATCCGCCAGGCCCAT  
GCAACATCAGCGCGAGAAAGTGGAAACAACACCCCTGAAGCAGATCGTGACCAAGCTGCAAGGCC  
CAGTTTCGGAACAAGACCATCTGTGTTCAAGCAGAGCAGCGCGCGGACCCCGAGATCTGTGAT  
GCACAGCTTCAACTGCGGCGCGGAGTTCTTCTACTGCAACAGCACCCAGCTGTTTAAACAGCAC  
CTGGAACAACACCATCGGCCCAACAACAACCAACGGCAACCATCAACCTGCGCTGCGCATCA  
AGCAGATCATCGGCGCGGCGCATGTACGCCCCCCCCATCCGCGGCCAGATCCGCTGCGAGCAGC  
AACATCAACCGCTGCTGTGACCCGCGCAGCGCGCAAGGAGATCAGCAAACACACCGAGAT  
CTTCCGCCCCGGCGCGCGCAGATGCGCGCAACTGCGCGCAGCGAGCTGTACAAGTACAAGT  
TGGTGAAGATCGAGCCCTGTGGCGTGGCCCCCAACGAAGGCCAAGCGCGCGTGGTGCAGCGC  
GAGAAGCGCGCCGTGACCTGGGCGGCATGTTCTTGGGCTTCTGGGCGCGCGCGCAGCAAC  
ATGGGCGCCCGCAGCCTGACCCGTGACCGTGACAGGCCCGCGCAGCTGCTGAGCGGCATCTGTGA  
GTCGGAACAACAACCTGTGTGCGGCCATCGAGGCCCAGCAGCACCTGCTGCACTGACCGTGT  
GGGGCATCAAGCAGCTGACGGCCCGCGTGTGCGCTGGAGCGCTACTGAGGACCAAGCAG  
CTGCTGGGCATCTGGGGCTGCAGCGCAAGCTGATCTGCACACCCGCGCTGCCCTGGAAGCGC  
AGCTGGAGCAACAAGAGCCTGGACCAAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCG  
CGAGATCGACAACATACACCAACCTGATCTACACCTGATCGAGGAGAGGCCAGAACGACGAG  
AGAAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCGAGCCTGTGGAACTGGTTCGAC  
ATCAGCAAGTGGCTGTGGTACATCAAGATCTTATCATGATCTGTGGGCGGCTGGTGGGCGTG  
CGCATCGTGTTCACCGTGTGAGCATCTGGAACCGCGTGGCCAGGGCTACAGCCCCCTGAGC  
TTCAGACCCGCTTCCCGCCCCCGCGGCCCGGACCGCCCCGAGGGCATCGAGGAGGAGGG  
CGCGAGCGCGCAACCGGACCGCAGCAGCCCTGTGTGACCGGCTGCTGGCCCTGATCTGGG  
ACGACCTGCGCAGCGCTGTGCTGTTCAGCTACCAACCGCTGCGCGACCTGATCTGATCGCGG  
CCGCACTGTGGAGCTGTGTGGCCGCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTG  
CTGCACTACTGGATCCAGGAGCTGAAGAAACAGCGCGTGTGAGCTGTTGACGCCATCGCCATC  
CCGCTGGCCGAGGGCAACCGACCGCATCATCGAGGTGGGCCAGCGCATCGGCCGCGCTTCT  
GCACATCCCCGCGCATCGGCCAGGCTTCGAGCGCGCCCTGCTGTAACCTGAG

FIG. 25

SEQ ID NO:23; VAL120-ILE201B; ILE424-ALA433

GAATTGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA  
GTCTTCCTTTCCGCCAGCGCGCTGGGAGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGGTG  
TGGAAAGGAGGCCACCAACCCCTGTTCTCGGCCACGCGACGCCAAGGCCATACGACAACCGAGGT  
GCACAACCGTGTGGGCCACCAACGCTGCGTCCACCGACCCCAACCCCAAGGAGATCGTGTCT  
GGAGAAACGTGACCGAGAACTTCAACATGTGGAAGAAACAACATGGTGGAGCAGATGCACGAG  
GACATCTATGACCTGTGGGACCAAGAGCTGAAGCCTGCGTGGCCGGCATCACCCAGGCCCTGC  
CCCAAGGTGAGCTTCGAGGCCATCCCATCCACTACTGCGCCCGCGCGGCTTCGCCATCTCTG  
AAGTGCAACGACAAGAAGTTCAACGGCAGCGGCCCTGCACCAACGTGAGGACCGTGCAGTG  
CACCCACGGCATCCGCCCGTGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGG  
AGGGCGTGGTGATCCGACGCGAGAATTCAACGACAACGCCAAGACCATCATCTGTGCAGCTG  
AAGGAGAGCGTGGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGAGCATCAACCAT  
CGGCCCGCGCGCGCTTCTACGCCACCGCGGACATCATCGGCGACATCCGCCAGGCCACTG  
CAACATCAGCGGCGAGAAGTGGAACAACACCTGAAGCAGATCGTGACCAAGCTGCAGGCC  
AGTTTGGCAACAAGACCATCTGTGTTCAAGCAGAGCAGCGCGCGGCAACCCGAGATCTGTGATG  
CACAGCTTCAACTGCGCGCGCGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACC  
TGGAACAACACCATCGGCCCAACAACCAACCGCACCATCAACCTGCGCTGCGCGCATCAA  
GCAGATCATCGCGCGCGCATGTACGCCCGCCCATCCGCGGCCAGATCCCGTGCAGCAGCA  
ACATCACCGGCGCTGCTGTGACCGCGCAGCGCGGCAAGGAGATCAGCAACACCAACCGAGATC  
TTCGCGCCCGCGCGCGCGCATGCGCGCAACATGGCGCAGCGAGCTGTACAAGTACAAGGT  
GGTGAAGATCGAGCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCGTGGTGAAGCGCG  
AGAAAGCGCGCGGTGACCTGGGCGCATGTTCTTGGGCTTCTGGGCGCGCGCGCAGCAACCA  
TGGGCGCCCGCAGCCTGACCCTGACCGTGACGGCCCGCGAGCTGCTGAAGCGGCATCTGTGAG  
CAGCAGAAACAACCTGTGCGCGCATCGAGGCCACAGCAGCACCTGCTGCAAGCTGAACGTGTG  
GGGCATCAAGCAGCTGCAGGCCCGCGTGTGCGCGTGGAAGCGCTACTGAAGGACAGCAGC  
TGCTGGGCATCTGGGCTGCAGCGCAAGCTGATCTGCACACCGCCGTGCCCTGGAACGCCA  
CTGGAGCAACAAGAGCCTGGACCGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGC  
GAGATCGACAATACACCAACCTGATCTACACCTGATCGAGGAGAGCCAGAACCCAGCAGGA  
GAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCGAGCCTGTGGAACCTGGTTCGACA  
TCAGCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCTGGGCGGCCCTGGTGGGCGTGC  
GCATCTGTGTTACCCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCGCTGAGCT  
TCCAGACCCGCTTCCCGCGCCCGCGCGGCCGACCGCCCGAGGGCATCGAGGAGGAGGGC  
GGCAGCGCGCAACCGCAGCGCAGCAGCCCTGTGGTGCACGGCTGCTGGCCCTGATCTGGGA  
CGACTCTGCGACGCTGTGCCGTGTTAGCTACCAACCGCTGCGGACCTGATCTGATCGCGCG  
CCGATCTGTGGAGCTGTGGGCGCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTGC  
TGCAGTACTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTGACGCGCATCGCCATC  
CGCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGGGCCAGCGCATCGGCGCGGCTTCTC  
GCACATCCCGCGCGCATCCGCCAGGGCTTCGAGCGCGCCTGCTGTAACCTGAG

FIG. 26

## SEQ ID NO:24 VAL120-THR202; ILE424-ALA433

GAATTCCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA  
GTCTTCGTTCGCCCCAGCGCCGTGGAGAGAAGCTGTGGGTGAACGCTGTACTACGGCGGTGCCCGTG  
TGGAAAGGAGGGCCACCACCACTGTTCTGCGCCAGCGACGCCAAGGCTTACGACACCGAGGT  
GCACAACGTGTGGGCCAACCAACGCTGCGTGCCCAACGCAACCCCAACCCCAAGGAGATCGTCT  
GGAGAAGCTGACCCGAGAACTTCAACATGTGGAGAACAACATGGTGGAGCAGATGACACGAG  
GACATCATCAGCCTGTGGGACCAGAGCTGAAGCCCTGCGTGGGCGGCGCCACCCAGGCGCTG  
CCCCAAGGTGAGCTTCGAGCCCATCCCATCCACTACTGCGGCCCGCGCCACCCAGGCGCTG  
GACCCCAACGCGATCCGCCCGTGGTGAGCAOCCAGCTGCTGCTGAACGCGACGCTGGCCGAG  
GAGGGCGTGGTGATCCGACGCGAGAACTTCAACGCAACGCCAAGACCATCATCGTGACGT  
GAAGGAGAGCGTGGAGATCAACTGCAACCGGCCCAACAACAACACCGCAAGAGCATCAACA  
TCGGCCCCGCGCGCGCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCATCT  
GCAACATCAGCGGCGAGAACTGGAAACAACACCCCTGAAGCAGATCGTGACCAAGCTGCAGGCC  
CAGTTCGGCAACAAGACCATCGTGTTCAGCAGAGCAGCGGCGGCGACCCGAGATCGTGTAT  
GCACAGCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCAGCTGTTCAACAGCAC  
CTGGAACAACAACCATGGGCCCAACAACAACAACCGCACCATACCTGCGCTGCGCGATCA  
AGCAGATCATCGGCGGCGCCATGTACGCCCGCCCATCCGCGGCGAGATCGCTGCAGCAGC  
AATCATACCGCGCTGCTGCTGACCCGCGGCGGCGCAAGGAGATCAGCAACAACACCGAGAT  
CTTCGCGCCCGCGCGCGGCGACATGCGCGACAACCTGGCGCAGCGAGCTGTACAAGTACAAGG  
TGGTGAAGATCGAGCCCTGGGCGTGGCCCCCAACCAAGGCCAAGCGCGCGTGGTGACGCGC  
GAGAAAGCGCGCTGACCCCTGGGCGCCATGTTCTGGGCTTCCTGGGCGCGCGCGGAGCAACC  
ATGGGCGCGCCGACGCTGACCTGACCGTGCAGGCCCGGCCAGCTGCTGAGCGGCATCGTGCA  
GCAGCAGAAACAACCTGCTGCGCGCCATCGAGGCCAGCAGCAACCTGCTGCACTGACCGTGT  
GGGGCATCAAGCAGCTGCAGGCCGCGTGTGCGCGTGGAGCGCTACCTGAAAGGACACGAG  
CTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCAACACCGCGCTGCCCTGGAAACGCC  
AGCTGGAGCAACAAGAGCCTGGACCAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCG  
CGAGATCGACAACATACACCAACCTGATCTACACCTGATCGAGGAGAGCCAGAACACGACGAG  
AGAAGAACGAGCAGGAGCTGCTGGAGCTGGAACAAGTGGGCCAGCCTGTGGAACCTGGTTCGAC  
ATCAGCAAGTGGCTGTGTACATCAAGATCTTATCATGATCGTGGGCGGCTGGTGGGCGCTG  
CGCATCGTGTTCACCGTGTGAGCATCGTGAAACCGCGTGGCGCAGGGCTACAGCCCGCTGAGC  
TTCAGACCCGCTTCCCCGCCCGCGCGCCCGACCGCGCGGAGGCGATCGAGGAGGAGGG  
CGGCGAGCGCGACCGGACCGCAGCAGCCCCCTGGTGACAGGCGCTGCTGGCCCTGATCTGGG  
ACGACCTGCGCAGCCTGTGCTGTTTACGACTACACCGCTGCGCGACCTGATCTGATCGCGG  
CCCGCTGGCTGGAGCTGTGTGGGCGCGCGGCTGGGAGGCCCTGAAGTACTGGGGCAACCTG  
CTGCACTACTGGATCCAGAGCTGAAGAACAGCGCGCGTGAGCTGTTTCAGCGCATCGCCATC  
GCGCTGGCCGAGGCGACCGACCGCATATCGAGGTGGCCAGCGCATCGCGCGCGCTTCCT  
GCACATCCCCCGCGCATCCGCCAGGCGCTCGAGCGCGCGCTGCTGAATCCAG

FIG. 27

## SEQ ID NO:25 VAL127-ASN195

GAATTCGCCACCATTGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA  
GTCTTCGTTTCGCCACGCGCGTGGAGAAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTG  
TGGAAAGGAGGCCACACCACCCCTGTTCTCGGCCACGCGACGCCAAAGGCTACGCAACCGAGGT  
GCACAACTGTGTGGGCCACCCACGCTCGGTGCCACCGACCCCAACCCCAAGGAGATCGTGCT  
GGAGAACTGTAACCGAGAACTTCAACATGTGGAAAGAAACATGGTGGAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAGAGCTGAAAGCCCTGCGTGAAGCTGACCCCGCTGTGGGTG  
GGGGCAGGGAACTCAACACCCAGCGTGATCACCCAGGCGTGGCCCAAGGTGAGCTTCGAGCC  
CATCCCCATCCACTACTGCGCCCCCGCGGCTTCGCCATCTGAAAGTGCAACGACAAAGAAAGT  
CAACGGCAGCGGCCCTGCACCAACGTGAGCACCCTGCAATGTCACCCACGGCATCCGCCCG  
TGGTGAGCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTGATCOGCAGC  
GAGAAGCTTCACCGACAACGCCAAGACCATCATCTGTCAGCTGAAAGGAGAGCGCTGGAGATCAA  
CTGCACCCGCCCAACAACAACACCCGCAAGAGCATCACCATCGGCCCGCGCGCTTCTA  
CGCCACCGGCGACATCATCGCGGACATCCGCCAGGCCACTGCAACATCAGCGGCGAGAAAGT  
GGAACAACACCCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCAAGTTCGGCAACAAGACCATC  
GTGTTCAAGCAGAGCAGCGCGCGGCGACCCGAGATCGTGATGCACAGCTTCAACTGCGGCGG  
CGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACCTGGAACAACACCATCGGCC  
CAACAACACCAACGGCAACATCAACCTGCCCCTGCCGCATCAAGCAGATCATCAACCCGCTGGC  
AGGAGGTGGGCAAGGCCATGTACGCCCGCCCATCCGCGGCGAGATCCGCTGCAGCAACAAC  
ATCACCCGCTGCTGCTGACCCGCGACGGCGCAAGGAGATCAGCAACAACCCAGAGATCTT  
CCGCCCGCGCGCGCGGACATGCGCGCAAACTGGCGCAGCGAGCTGTACAAGTACAAGGTGG  
TGAAAGATCAGAGCCCTGGGCGTGGCCCCCAAGGCCAAGCGCGCGTGGTGACGCGGAG  
AAGCGCGCGGTGACCTGGGCGCCATGTTCTGGGCTTCTGGGCGCGCGCGGAGCAACCATG  
GGCGCCCGCAGCCTGACCTGACCGTGCAGGCGCGCCAGCTGCTGAGCGGCATCTGTCAGCA  
GCAGAAACAACCTGCTGCGCGCCATCGAGGCCAGCAGCACTGCTGCACTGACCTGTGGG  
GCTCAACGACGCTGCAGGCCCGCGTGTGCGCCGTGGAGCGCTACCTGAAAGGACCAAGCACTG  
CTGGGCATCTGGGCGTGCAGCGGCAAGCTGATCTGCACCAACCGCGCTGCCCTGGAAGCCAG  
CTGGAGCAACAAGAGCCTGGACAGATCTGGAACAACATGACCTGGATGAGTGGAGGCGCG  
AGATCGCAACTACACCAACCTGATCTACACCTGATCGAGGAGAGGCCAGAACCCAGCAGGAG  
AAGAACGAGCAGGAGCTGTGGAGCTGAGCAAGTGGGCCAGCGCTGTGGAATCTGGTTGACAT  
CAGCAAGTGGCTGTGGTACATCAAGATCTTATCATGATCGTGGCGCGCTGGTGGGCGCTGCG  
CATCGTGTTCACCGTGTGAGCATCTGTGAACCGCGTGCGCCAGGGCTACAGCCCGCTGAGCTT  
CCAGACCCGCTTCCCGCCCCCGCGGCCCGACCGCCCGAGGGCATCGAGGAGGAGGGCG  
GCGAGCGCGACCGCGACCGCAGCAGCCCGCTGTTGACCGGCTGCTGGCCCTGATCTGGGAC  
GACCTGCGCAGCCTGTGCTGTTTCACTACACCGCTGCGGACCTGATCTGATCGCGGCC  
CGCATCTGTGGAGCTGTGGGCGCGCGGCGTGGGAGGCCCTGAAGTACTGGGGCAACCTGCT  
GCACTATCGGATCCAGGAGCTGAAGAACAGCGCGCTGAGCGCTGTCGACGCCATCGCCATCG  
CGTGGCCGAGGGCAGCGACCGCATATCGAGGTGGCCACGCGCATCGGCCCGCGCTTCTCGT  
ACATCCCCCGCGCATCCGCCAGGGCTTCGAGCGCGCGCTGCTGTAACTCGAG

FIG. 28

SEQ ID NO:26 VAL127-ASN195; ARG426-GLY431

GAATTGCGCCACCATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCA  
GTCTTCGTTCGCCACGCGCGTGGAGAAAGCTGTGGGTGACCGTGTACTACGGCGTGCCTGGT  
TGGAAAGGAGGGCCACCAACCCCTGTTCTGCGCCAGCGACGCCAAGGCCTACGACACCGAGGT  
GCACAACGCTGTGGGCCACCCACGCTCGTGGCCACCGACCCCAACCCCAAGGAGATCGTGCT  
GGAGAACGTGACCGGAACTTCAACATGTGGAAGAAACAACATGTGTGGAAGCAGATGCACGAG  
GACATCATCAGCCTGTGGGACCAGAGCTGAAGCCCTGCGTGAAGCTGACCCCTGTGGGTG  
GGGGCAGGGGAAGCTCAACACCCAGCGTGATCAACCGGCTGCGCCCAAGGTGAGCTTCGAGCC  
CATCCCATCCACTACTGCGCCCCCGCGGCTTGCCATCTGAACTGCAACGACAAGAAGTT  
CAACGGCACCGGCCCTGCACCAACGCTGAGCACCGTGCAGTGCACCCACGGGATCCGCCCG  
TGGTGAGCACCCAGCTGCTGTGTAACGGCAGCTGCGCGGAGGAGGGGTGGTGATCCGCAGC  
GAGAACTTCACCGACAACGCCAAGACCATCATGTGTCAGCTGAAGGAGAGCGTGGAGATCAA  
CTGCACCCGCCCAACAACAACCCGCAAGAGCATCACCATCGGCCCGCGCGCTCTTA  
CGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCACTGCAACATCAGCGGCGAGAAAT  
GGAAACAACCCCTGAAGCAGATCGTGACCAAGCTGCAGGCCCACTTCGGCAACAAGACCATC  
GTGTTCAAGCAGAGCAGCGGCGGCGACCCGAGATCGTGATGCACAGCTTCAACTGCGGCGG  
CGAGTTCTTCTACTGCAACAGCACCCAGCTGTTCAACAGCACTTGAACAACAACATCGGCC  
CAACAACAACCAACGGCACCATCAACCTGCGCTGCGCATCAAGCAGATCATCAACCGCGCG  
GCGGCAAGGCCATGTACGCCCCCCCATCCGCGGCGCAGATCCGCTGCAGCAGCAACATCACC  
GGCTGTGCTGACCCGCGACGCGGCGCAAGGAGATCAGCAACACCCAGCAGATCTTCGCGCC  
CGGGGCGGCGACATGCGCGCAACAAGCTGCGCGCAGCGAGCTGTACAAGTACAAGGTGGTGAAG  
ATCGAGCCCTGGGCGTGGCCCCACCAAGGCCAAGCGCCGCTGGTGCAAGCGCGAGAAGCG  
CGCGGTGACCTGGGCGCCATGTTCTCGGCTTCTGGGCGCGCGCGGCGACCAATGGGCGC  
CGCAGCTGACCTGACCTGACCTGACCGCGGCGACCGCCGCACTGCTGAGCGGCATCGTGCAGCAGCA  
ACAACCTGCTGCGCGCCATCGAGGCCAGCAGCCTGCTGCAGCTGACCGTGTGGGGCATC  
AAGCAGCTGCAGGCCCGCGTGTGCGCGTGAGCGCTACCTGAAGGACAGCAGCTGCTGGG  
CATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCAACCGCGTGCCTGGAAACGCCAGCTGGA  
GCAACAAGAGCTGGACAGATCTGGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATC  
GACAACACCAACAACCTGATCTACACCTGATCGAGGAGAGCCAGAAACCAAGAGAGAAGAA  
CGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCGAGCCTGTGGAACCTGTTTCGACATCAGCA  
AGTGGCTGTGGTACATCAAGATCTTCAATCATGATCGTGGGCGCGCTGGTGGGCTCGCGCATCG  
TGTTACCGTGTGAGCATCGTGAACCGCGTGGCGCAGGGCTACAGCCCTGAGCTTCCAGA  
CCCGCTTCCCCGCGCCGCGCGCCGACCGCCGAGGGCATCGAGGAGGAGGGCGCGAG  
CGCAGCGCGACCGCAGCAGCGCCCTGGTGACCGGCTGCTGGCCCTGATCTGGGAGCAGCCTG  
CTGGAGCTGTGGGCGCGCGGCTGGGAGGCCCTGAAGTACTGGGGAACCTGCTGCAGTA  
CTGGATCCAGGAGCTGAAGAACAGCGCCGTGAGCCTGTTTCGACGCCATCGCCATCGCCGTGG  
CCGAGGGCACCGACCGCATCATCAGGTGGCCAGCGCATCGGCCCGCGCTTCTGCACATCC  
CCCGCGCATCCGCCAGGGCTTCGAGCGCGCCTGCTGTAACTCGAG

FIG. 29



## SEQUENCE LISTING

&lt;110&gt; Chiron Corporation

&lt;120&gt; MODIFIED HIV ENV POLYPEPTIDES

&lt;130&gt; 1605.100

&lt;140&gt;

&lt;141&gt;

&lt;160&gt; 26

&lt;170&gt; PatentIn Ver. 2.0

&lt;210&gt; 1

&lt;211&gt; 856

&lt;212&gt; PRT

&lt;213&gt; Human immunodeficiency virus

&lt;400&gt; 1

```

Met Arg Val Lys Glu Lys Tyr Gln His Leu Trp Arg Trp Gly Trp Arg
 1             5             10             15

Trp Gly Thr Met Leu Leu Gly Met Leu Met Ile Cys Ser Ala Thr Glu
          20             25             30

Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala
 35             40             45

Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu
 50             55             60

Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn
 65             70             75             80

Pro Gln Glu Val Val Leu Val Asn Val Thr Glu Asn Phe Asn Met Trp
          85             90             95

Lys Asn Asp Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp
      100             105             110

Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Ser
      115             120             125

Leu Lys Cys Thr Asp Leu Lys Asn Asp Thr Asn Thr Asn Ser Ser Ser
      130             135             140

Gly Arg Met Ile Met Glu Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn
      145             150             155             160

Ile Ser Thr Ser Ile Arg Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe
      165             170             175

Tyr Lys Leu Asp Ile Ile Pro Ile Asp Asn Asp Thr Thr Ser Tyr Lys
      180             185             190

```

Leu Thr Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val  
 195 200 205  
 Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala  
 210 215 220  
 Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr  
 225 230 235 240  
 Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser  
 245 250 255  
 Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Val Val Ile  
 260 265 270  
 Arg Ser Val Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu  
 275 280 285  
 Asn Thr Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg  
 290 295 300  
 Lys Arg Ile Arg Ile Gln Arg Gly Pro Gly Arg Ala Phe Val Thr Ile  
 305 310 315 320  
 Gly Lys Ile Gly Asn Met Arg Gln Ala His Cys Asn Ile Ser Arg Ala  
 325 330 335  
 Lys Trp Asn Asn Thr Leu Lys Gln Ile Ala Ser Lys Leu Arg Glu Gln  
 340 345 350  
 Phe Gly Asn Asn Lys Thr Ile Ile Phe Lys Gln Ser Ser Gly Gly Asp  
 355 360 365  
 Pro Glu Ile Val Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr  
 370 375 380  
 Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Phe Asn Ser Thr Trp  
 385 390 395 400  
 Ser Thr Glu Gly Ser Asn Asn Thr Glu Gly Ser Asp Thr Ile Thr Leu  
 405 410 415  
 Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Lys Val Gly Lys  
 420 425 430  
 Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg Cys Ser Ser Asn  
 435 440 445  
 Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Ser Asn Asn Glu  
 450 455 460  
 Ser Glu Ile Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg  
 465 470 475 480  
 Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val  
 485 490 495  
 Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg Ala  
 500 505 510

Val Gly Ile Gly Ala Leu Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser  
 515 520 525  
 Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val Gln Ala Arg Gln Leu  
 530 535 540  
 Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu  
 545 550 555 560  
 Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu  
 565 570 575  
 Gln Ala Arg Ile Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu  
 580 585 590  
 Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val  
 595 600 605  
 Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu Glu Gln Ile Trp Asn  
 610 615 620  
 His Thr Thr Trp Met Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser  
 625 630 635 640  
 Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn  
 645 650 655  
 Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp  
 660 665 670  
 Phe Asn Ile Thr Asn Trp Leu Trp Tyr Ile Lys Leu Phe Ile Met Ile  
 675 680 685  
 Val Gly Gly Leu Val Gly Leu Arg Ile Val Phe Ala Val Leu Ser Ile  
 690 695 700  
 Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr His  
 705 710 715 720  
 Leu Pro Thr Pro Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu Glu  
 725 730 735  
 Gly Gly Glu Arg Asp Arg Asp Arg Ser Ile Arg Leu Val Asn Gly Ser  
 740 745 750  
 Leu Ala Leu Ile Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr  
 755 760 765  
 His Arg Leu Arg Asp Leu Leu Leu Ile Val Thr Arg Ile Val Glu Leu  
 770 775 780  
 Leu Gly Arg Arg Gly Trp Glu Ala Leu Lys Tyr Trp Trp Asn Leu Leu  
 785 790 795 800  
 Gln Tyr Trp Ser Gln Glu Leu Lys Asn Ser Ala Val Ser Leu Leu Asn  
 805 810 815  
 Ala Thr Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Val Ile Glu Val  
 820 825 830

Val Gln Gly Ala Cys Arg Ala Ile Arg His Ile Pro Arg Arg Ile Arg  
835 840 845

Gln Gly Leu Glu Arg Ile Leu Leu  
850 855

<210> 2

<211> 847

<212> PRT

<213> Human immunodeficiency virus

<400> 2

Met Arg Val Lys Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Gly  
1 5 10 15

Gly Thr Leu Leu Leu Gly Met Leu Met Ile Cys Ser Ala Val Glu Lys  
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr  
35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val  
50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro  
65 70 75 80

Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys  
85 90 95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp  
100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu  
115 120 125

His Cys Thr Asn Leu Lys Asn Ala Thr Asn Thr Lys Ser Ser Asn Trp  
130 135 140

Lys Glu Met Asp Arg Gly Glu Ile Lys Asn Cys Ser Phe Lys Val Thr  
145 150 155 160

Thr Ser Ile Arg Asn Lys Met Gln Lys Glu Tyr Ala Leu Phe Tyr Lys  
165 170 175

Leu Asp Val Val Pro Ile Asp Asn Asp Asn Thr Ser Tyr Lys Leu Ile  
180 185 190

Asn Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe  
195 200 205

Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu  
210 215 220

Lys Cys Asn Asp Lys Lys Phe Asn Gly Ser Gly Pro Cys Thr Asn Val  
225 230 235 240

Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln  
 245 250 255  
 Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Gly Val Val Ile Arg Ser  
 260 265 270  
 Glu Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu Lys Glu  
 275 280 285  
 Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser  
 290 295 300  
 Ile Thr Ile Gly Pro Gly Arg Ala Phe Tyr Ala Thr Gly Asp Ile Ile  
 305 310 315 320  
 Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser Gly Glu Lys Trp Asn  
 325 330 335  
 Asn Thr Leu Lys Gln Ile Val Thr Lys Leu Gln Ala Gln Phe Gly Asn  
 340 345 350  
 Lys Thr Ile Val Phe Lys Gln Ser Ser Gly Gly Asp Pro Glu Ile Val  
 355 360 365  
 Met His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Ser Thr  
 370 375 380  
 Gln Leu Phe Asn Ser Thr Trp Asn Asn Thr Ile Gly Pro Asn Asn Thr  
 385 390 395 400  
 Asn Gly Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Arg  
 405 410 415  
 Trp Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln  
 420 425 430  
 Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly  
 435 440 445  
 Gly Lys Glu Ile Ser Asn Thr Thr Glu Ile Phe Arg Pro Gly Gly Gly  
 450 455 460  
 Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val  
 465 470 475 480  
 Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val  
 485 490 495  
 Val Gln Arg Glu Lys Arg Ala Val Thr Leu Gly Ala Met Phe Leu Gly  
 500 505 510  
 Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Arg Ser Leu Thr Leu  
 515 520 525  
 Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Asn  
 530 535 540  
 Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr  
 545 550 555 560

Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg  
 565 570 575  
 Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys  
 580 585 590  
 Leu Ile Cys Thr Thr Ala Val Pro Trp Asn Ala Ser Trp Ser Asn Lys  
 595 600 605  
 Ser Leu Asp Gln Ile Trp Asn Asn Met Thr Trp Met Glu Trp Glu Arg  
 610 615 620  
 Glu Ile Asp Asn Tyr Thr Asn Leu Ile Tyr Thr Leu Ile Glu Glu Ser  
 625 630 635 640  
 Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys  
 645 650 655  
 Trp Ala Ser Leu Trp Asn Trp Phe Asp Ile Ser Lys Trp Leu Trp Tyr  
 660 665 670  
 Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Val Gly Leu Arg Ile  
 675 680 685  
 Val Phe Thr Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser  
 690 695 700  
 Pro Leu Ser Phe Gln Thr Arg Phe Pro Ala Pro Arg Gly Pro Asp Arg  
 705 710 715 720  
 Pro Glu Gly Ile Glu Glu Glu Gly Gly Glu Arg Asp Arg Asp Arg Ser  
 725 730 735  
 Ser Pro Leu Val His Gly Leu Leu Ala Leu Ile Trp Asp Asp Leu Arg  
 740 745 750  
 Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Leu Ile Leu Ile  
 755 760 765  
 Ala Ala Arg Ile Val Glu Leu Leu Gly Arg Arg Gly Trp Glu Ala Leu  
 770 775 780  
 Lys Tyr Trp Gly Asn Leu Leu Gln Tyr Trp Ile Gln Glu Leu Lys Asn  
 785 790 795 800  
 Ser Ala Val Ser Leu Phe Asp Ala Ile Ala Ile Ala Val Ala Glu Gly  
 805 810 815  
 Thr Asp Arg Ile Ile Glu Val Ala Gln Arg Ile Gly Arg Ala Phe Leu  
 820 825 830  
 His Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Arg Ala Leu Leu  
 835 840 845

<210> 3  
 <211> 2310  
 <212> DNA  
 <213> Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Val120-Ala204

&lt;400&gt; 3

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgtgtgt gctgtgtgga 60
gcagctcttcg ttctgcccag cgcgtggagc aagctgtggg tgaccgtgta ctacggcgctg 120
cccggtgtgga aggaaggccac caccaccctg ttctgcccga gcgacgcccga ggccctacgac 180
accgaggtgac acaagctgtg ggcacccacc gctctgctgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggcacatcat cagcctgtgg gaccagagcc tgaagccctg cgtggggcgc 360
ggcgctctgcc ccaaggtgag ctctgagccc atccccatcc actactgcgc ccccgcgagc 420
ttcgccatcc tgaagtgcac cgacaagaag ttcaacggca gcggcccctg caccacaagt 480
agcaccgtgc agtgacccca cggcatccgc cccgtgggtg gacccagat gctgtgtgac 540
ggcagccctg cggagggagg cgtggtgatc cgcagcgaga acttcaccca caacgccaac 600
accatcatcg tgcagctgaa ggagagcgtg gagatcaact gcaccgcgcc caacaacaa 660
accgcgaaga gcatcaccat cggcccgcgc cgcgcttctc acgcacccgc cgacatcatc 720
ggcgacatcc gccagggccc ctgcaacatc agcggcgaga agtggaaaca cccctggaag 780
cagatcgtga ccaagctgca gcccagctt ggcaacaaga ccactcgtgt caagcagagc 840
agcggcgagg accccagat cgtgatccac agcttcaact gcggcgcgca gttcttctac 900
tgcaacagca cccagctgtt caacagcacc tggaaacaaa ccacggccc caacaacacc 960
aagggcacca tcacctctgc ctgcccgcac aagcagatca tcaaccgctg gcaggaggtg 1020
ggcaaggcca tgtacgcccc ccccatccgc ggcacatccc gctgcagcag caacataccc 1080
ggcctctctgc tgaccgcgga cggcgccaag gagatcagca acaccacgga gatcttccgc 1140
ccccggcgcg cgcacatcgc cgacaactgg cgcagcgagc tgtacaagta caaggtctgt 1200
aagatcgagc cctctgggctg ggccccccac aaggccaagc gccgctggtg gcagcgcgct 1260
aaacgcgcgc tgaccctggg cgccatgttc ctgggtcttc tggcgcgccg cgcgacgacc 1320
atggggcgccc gcagcctgac cctgacctgc caggcccgcg agctgctgag cgcactcgtg 1380
cagcagcaga caaacctgct gcgcgccatc gaggcctgct agcacctgct gcagctgacc 1440
gtgtgtgggca tcaagcagct gcagggccgc gtgctggcgc tggagcgcta cctgaaggac 1500
cagcagctgc tggcatctg gggctgcagc ggcaagctga tctgcacca cgcgctgccc 1560
tggaaagcca gctggagcaa caagagcctg gaccagatct ggaacaacat gacctggatg 1620
gagtgggagc gctgagatga caactacacc aacctgatct acacctgat cgaggagagc 1680
gaacaaacagc aggaagaaga cgacaggagc ctgctggagc tggacaagt ggcagcctg 1740
tggaaactggt tcgacatcag caagtggctg tggatcatca agatcttcat catgatcctg 1800
ggcgccctggt tggccctgcg catcgtgttc accgtgctga gcatcgtgaa ccgctgtgcg 1860
cagggtctaca gccctctgag ctctcagacc cgcttcccgc ccccgcgcgc ccccgagcgc 1920
cccgaaggga tcgaggagga gggcgcgagc gcgcacgcgc accgcagcgc cccctgtgtg 1980
cacggcctgc tggccctgat ctgggacgac ctgcgcagcc tbtgctgtgt cagctacacc 2040
cgccctgcgc acctgatcct gatcgccgccc cgcactgtgg agctgctggg ccgcccgcgc 2100
tgggagggccc tgaagactgc gggcaacctg ctgcagctat ggaatcagga gctgaagaac 2160
agcgcgctga gctctgttga cgcacatccc atcgctgtgg ccgagggcac catggtggag 2220
atcgaggtgg cccagcgcat cggcgcgccc ttctctgaca tccccgcgcg catccgcgac 2280
ggcttcgagc ggcctctgct gtaactcgag

```

&lt;210&gt; 4

&lt;211&gt; 2316

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Val120-Ile201

&lt;400&gt; 4

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgtgtgt gctgtgtgga 60
gcagctcttcg ttctgcccag cgcgtggagc aagctgtggg tgaccgtgta ctacggcgctg 120
cccggtgtgga aggaaggccac caccaccctg ttctgcccga gcgacgcccga ggccctacgac 180
accgaggtgac acaagctgtg ggcacccacc gctctgctgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggcacatcat cagcctgtgg gaccagagcc tgaagccctg cgtggggcgc 360

```

```

atcaccacag cctgccccaa ggtgagcttc gagccatcc ccatccacta ctgcgcccc 420
gcgcgtctcg ccatcctgaa gtgcaacgac acggcagcgg cccctgcacc 480
acacgtgagca ccgtgcagtg caccacccgg atccgcgccg tgggtgagcac ccagctgtctg 540
ctgaacggca gccctggcga ggaagggcgtg gtgatccgca gcgagaacct caccgcacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgcccacaac 660
aacaacaccc gcaagagcat caccatccgg cccgcgccgg ccttctacgc caccggcgac 720
atcatcgggc acatccgcca ggcccactgc gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagagcagcg gcggcgaccc cgagatcgtg atgcacagct tcaactcgcg cggcgagttc 900
ttctactgca acagcaccca gctgttcaac agcactcgga acaacaccat cggccccaac 960
aacaccaacg gcaccaatcac cctgcctcgc cgcatacaca agatcatcaa cctctggcag 1020
gaggtgggca agggcatgta cgcgccccc atccgcggcc agatccgctg cagcagcaac 1080
atcacccggc tgcgtcgtgac ccgcgacggc ggcaaggaga tcagcaaac caccgagatc 1140
ttccgccccg gcggcgcgca catgcgcgac aactggcgca gcgagctgta caagtacaag 1200
gtggtgaaag tcgagccctt gggcgtggcc cccaccaagg ccaagcgccg cgtggtgcag 1260
cgcgagaagc gcgcgctgac cctggggccc atgttccctg gcttccctgg cgcgcgcggc 1320
agcacaatgg gcgcgcgacg cctgacccct acocgtgcag cccgcagctg cctgagcggc 1380
atcgtgcagc agcaagaacaa cctgctgcgc gccatcgagg cccagcagca cctgctgcag 1440
ctgacccgtg ggggcatcaa gcagctgcag gcccgctgc tggccgtgga gcgctacctg 1500
aaggaccagc agctcgtggg catctggggc tgcaggggca agctgatctg caccacggcc 1560
gtgcctcgga accgcagctg gacgaacaag agcctggacc agatctggaa caacatgacc 1620
tggatggagt gggagcgcgca gatcgacaac tacaccaacc tgatctacac cctgctcagc 1680
gagagccaga accagcagga gaagaaacag caggagctgc tggagctgga caagtgggcc 1740
agcctgtgga actggttcga catcagcaag tggctgtggt acatcaagat ctctcatctg 1800
atcgtggggc gccctggtgg cctgcgcctc gtgttccacc tgcagcatc cgtgaaccgc 1860
gtgcgcaggc gctacagccc cctgagcttc cagacccccc tccccgccc 1920
gacgcgcccg agggcatcga ggaggaggcc ggcgagcgcc accgcgacgg cagcagcccc 1980
ctggtgcacg cctgctggc cctgatctgg gcgacccctg cgcgctcagc cctgttcagc 2040
taccacccgc tgcgcgaccc gatcctgac gccgcgccga tcgtggagct cgtggcgccg 2100
cgcgcctggg agcgcctgaa gtactggggc aacctgctgc agtactggat cagggagctg 2160
aagaacagcg ccgtgagcct gttcgacgcc atccgcatcc cctggtcgga gggcaccgac 2220
cgcctcatcg aggtggccca gcgcactcgc cgcgccttcc tgcacatccc cgcgcgctc 2280
cgccagggct tcgagcgcgc cctgctgtaa ctcgag 2316

```

<210> 5  
 <211> 2322  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Val120-Ile201B

```

<400> 5
gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgtgctg gctgtgtgga 60
cgactcttcg ttcgcccgag cgcgctggag aagctgtggg tgaccgtgta ctacggcggtg 120
cccgcttgga agagggccac caccaccctg tctctgcgca gcgacgccaa ggctcagac 180
accgagtgcc acaacgtgtg ggccaacccc gectgcgtgc ccacgacccc caacccccag 240
gagatcgtgc tggagaacct gaccgagaac ttcaacatgt ggaagacaaa catgtggag 300
cagatgcacg aggaatcatc cagcctgtgg gaccagagccc tgaagccctg cgtgcccgcc 360
atcaccacgg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgcccc 420
gcgcgtctcg ccatcctgaa gtgcaacgac aagaagttca acggcagcgg cctctgaccc 480
aacgtgagca ccgtgcagtg caccacccgg atccgcgccg tgggtgagcac ccagctgtctg 540
ctgaagcaga gccctggcga ggaagggcgtg gtgatccgca gcgagaacct caccgacaac 600
gccaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgcccacaac 660
aacaacaccc gcaagagcat caccatccgc cccgcgcggc ccttctacgc caccggcgac 720
atcatcgcgg acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagagcagcg gcggcgaccc cgagatcgtg atgcacagct tcaactcgcg cggcgagttc 900
ttctactgca acagcaccca gctgttcaac agcactcgga acaacaccat cggccccaac 960

```



```

aacaccaacg gcaccatcac cctgccctgc cgcataaagc agatcatcaa ccgtggccag 1020
gaggtggggc aggccatgta cgcgccccc atccggcgcc agatccgctg cagcagcaac 1080
atcacccggc ttctgtgctac cgcgcagcgg ggcaaggaga tcagcaaacac caccgagatc 1140
ttccgcccgc gggcgggcga catcgccgac aactggcgca gcgagctgta caagtacaag 1200
tggtgaaaga tggagccctt gggcgtggcc cccaccaaag ccaagcgccg cgtgtacgag 1260
cgcgagaagc ggcgcgtgac cctggggcgc atgttccctg gcttcctggg ccgcgcggcg 1320
agcaccatgg ggcgcccgag cctgaccctg accgtgcagg ccgcgcagct cgtgagcgcg 1380
atcgtgcagc agcagaacaa cctcgtgcgc gccatcgagg ccacgacga cctgctgcag 1440
ctgaccctgt ggggcacatc caagcgtgcg gcccgctgdc tggcgtgga gcgtacctg 1500
atggaccagc agctcgtggg catctggggc tgcagcgcca agctcatctg caccaccgac 1560
tgcccttggg acgcacagct gagcaacaag agcctggacc agatctggaa caactgacac 1620
tggatggagt gggagcgcca gatcgacaac tacaccaaac tgatctacac cctgatcgag 1680
gagagccaga accagcagga gaagaacgag caagagctgc tggagctgga caagtgggac 1740
agccttggga actggttcca catcagcaag tggctgtgtg acatcaagat ttctcatgat 1800
atcgtggggc gctcgtggg cctgcgcctc gtgttccacc tctgagcat cgtgaaccgc 1860
gtgcgccagg gctacagccc cctgagcttc cagaccctgc tccccccccc ccgcggcccc 1920
gaccgccccc agggcatcga ggaggaggcg ggagagcgcg accgcgaccg cagctacccc 1980
ctggtgcagc gctcgtggc cctgatctgg gacgacctgc gacgctctgt cctgttccgc 2040
taccaccccc tggcgacct gatcctgate gcgcgccgca tctgtggagt cgtgagctg 2100
cgcggtgggg agggcctgaa tgaactgggg aacctgctgc agtactggat ccaggagctg 2160
aagaacagcg ccgtgagcct gtctgacgac atcgccatcg ccgtggcgga gggcaccagc 2220
cgcatcatcg aggtggccca gcgcacggc cgcgctcttc tgcacatccc ccgcgcgcat 2280
cgccagggct tggagcgcg cctcgtgtaa ctcgagcgtg ct 2322

```

&lt;210&gt; 6

&lt;211&gt; 2328

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Lys121-Val200

&lt;400&gt; 6

```

gaattcgcca ccatggatgc aatgaagaga gggctctgtg gtgtgctgtg gctgtgtgga 60
cgagttcttg ttctgcccag cgcggtggag aagctctggg tgacctgtga ctacggcgtg 120
ccggtgtgga aggaaggccac caccaccctg ttctgcgcca gcgacgccaa ggctctagac 180
accgaggttg acaacgtgtg ggccaccacc gcttgcgtgc ccaccgacc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaga ttcaacatgt ggaagaacaa ctgtgtggag 300
cagatgcacg aggcacatcat cagcctgtgg gaccagagcg tgaagccctg cgtggaagcc 360
cccgatgata cccagccctg ccccaagggt agcttcgagc ccactccccat ccaactatgc 420
gcccgcgcgc gcttgcacct cctgaagctc aacgacaaga agttcaacgg cagcggcccc 480
tgacacaaag tgagcaccgt gcagtgcacc cagccgctac gccctgtggg gagcaccag 540
ctgctgtgta agggcagcct ggccgaggag ggctgtgtga tccgcagcga gaacttcaac 600
gacaaagcca agaccatcat cgtgcagctg aaggagagcg tggagatcaa ctgcaccgcg 660
cccaacaaca acaccgcgaa gagcatcacc atcggccccc gcgcgcctt ctacgccacc 720
ggcgacatca tcggcgacat ccgccagggc cactgcacaa tcagcggcga gaagtggaa 780
aacaccctga agcagatcgt gaccaagctg caggccagct tcggcaacaa gacctcgtg 840
ttcaagcaga gcagcggcgg gacccccagc atcgtgatgc acagttcaa ctgcggcgcg 900
gagttcttct actgcaacag caccacagct ttcaacagca cctggaacaa caccatcggc 960
cccaacaaca ccaacggcac catcacctgt ccttcgcgca tcaagcagat catcaaccgc 1020
tggcaggagg tggcgaaagg catgtacgac ccccctatcc gcggccagat ccgctgcgac 1080
agcaacatca ccggcctgct gctgacccgc gacggcgcca aggagatcag caacaccaa 1140
gagatcttcc gcccccgcgg cggcgacatg cgcgacaact ggcgacgca gctgtacgag 1200
tacaaggttg tgaagatcga gccctctggc gtggccccca ccaaggccaa gcgcgcgctg 1260
gtgcagcgcg agaagcgcg cgtgacccct ggccgcatgt tctctggctt cgtggcgcc 1320
gcggcgagca ccatggcgcc cgcgagcctg accctgacgt tgcagccctg ccagctgctg 1380
agcggcatcg tgcagcaga gaacaacctg ctgcgcgcca tcgagggcca gcagcactg 1440
ctgcagctcg cctgctggg catcaagcag cgtgactggc cgtgagcgcc 1500
taacctgaag accagcagct gctgggacac tggggctgca gcggcaagct gatctgcacc 1560

```

```

accgcgcgtgc cctggaacgc cagctggagc aacaagagcc tggaccagat ctggaacaac 1620
atgacctgga tggngtggga gcgcgagatc gacaaactaca ccaacctgat ctacacctg 1680
atcgaggaga gccagaacca gcaggagaa aacgagcagg agctgctgga gctggacaag 1740
tggggcagcc tgtggaactg gttcgacatc agcaagtggc tgtggtacat caagatcttc 1800
atcatgatcg tggcgggcct ggtgggcctg cgcacgtgtg tcaccgctgt gacgatctgt 1860
aacgcgtgic gccagggcta cagccccctg agcttccaga cccgcttccc cggccccgcg 1920
ggccccagcc gccccgaggg catcgaggag gaggcgcgcg agcgcgaccg cgaccgcagc 1980
agccccctgg tgcacggcct gctggccctg atctgggagc acctgcccag cctgtgctgt 2040
ttcagctacc agcgccctgc cgacctgac ctgatcgccg cccgatcgct gtagctgtgt 2100
ggcgcgcgcg gctgggagcg cctgaagtag tggggcaacc tgctgagta ctggatccag 2160
tgctgaaga acagcgccgt gagcctgttc gacggcatcg ccatcgccgt cggcgaggcg 2220
accgacgcga tcctcgaggt gggccagcgc atcgccgcgc ccttctcgca catccccgcg 2280
cgatccgcgc agggcttcga gcgcgcctcg ctgtaactcg agcgctgtc 2328

```

&lt;210&gt; 7

&lt;211&gt; 2334

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Leul22-Ser199

&lt;400&gt; 7

```

gaattcgcca ccatggatgc aatgaagaga gggctctgtg gtgtgtgtgt gctgtgtgga 60
gcagctcttc tttcgccacg cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccggtgaga aggagggcac caccaccctg tcttcgcgca gcgacgccaa ggcctacgac 180
accaggttgc acaacgtgtg ggccaccacc gcttgcgtgc ccacgcagcc caacccccag 240
cagatcgtgc tggagaacgt gaccgagaac ctcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggaacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
ggcaacagcg tgatcaccca ggccctgccc aagtgagct tcgagcccat ccccatccac 420
tactgcgccc cgcgcggctt cgccatcctg aagtgcacg acaagaagt caacggcagc 480
ggccccgca ccaacgtgag caccgtgcag tgcacccac gcatccgccc cgtggtgagc 540
accagctgic tgctgaacgg cagcctggcc gaggaggggc tggatgacc cagcgagaac 600
ttcacgcaga acgccaagac catcatcgtg cagctgaagc agagcgtgga gatcaactgc 660
accgccccca acaacaacac ccgcaagagc ataccatcg gcccgcgcc cgccctctac 720
ggcaaccgcg acatcatcgg cgacatccgc caggcccaat gcaacatcg cggcgagaag 780
tggacaacca cctgaagca gatcgtgacc aagctgcagg cccagttcgg caacaagacc 840
atcgtgttca agcagagcag cggcgggcag cccgagatcg tgatgcacag cttcaactgc 900
ggcgccgagt tcttctactg caacagcacc cagctgttca acagcacctg gaacaacacc 960
atcgccccca acaaacccaa cggcaccatc accctgcccc gcgcatcaa cgagatcttc 1020
aacccgtggc aggaggtggg caaggccatg ccatccggcg ccatcgccgc cagatccgcg 1080
tgacagcaga acatcacccg cctgctgtgt acccgcgagc gggcgcaagg gatcagaac 1140
accacgaga tcttcgcccc cggcgggcgc gacatgcgcg acaactggcg ccaactggcg 1200
tacaagtaca agtggtgaa gatcgagccc ctgggctgtg ccccaacca ggccaagcgc 1260
cgctgtgtgc agcgcgagaa gcgcgcgtg accctgggcg ccatgttccc gggcttccc 1320
ggcgccgcgc gcagcaccat gggcgcccgc agcctgaccc tgacgtgca ggccccgac 1380
ctcctgagcg gcatcgtgca gcagcagaac aacctgctgc gcgcatcg cgccccagc 1440
cacctgtgic agctgacagt gtggggatc aagcagctgc aggcgccgct gctggcgctg 1500
gagcgtctacc tgaaggacca gcagctgctg ggcattctgg gctgcagcgc caagctgac 1560
tgcaacaccc cgtgcccctg gaacggcagc tggagcaaca agagcctgga ccagatctg 1620
aacaacatga cctgtgagg gttgggagcg gagatcgaca actacccaa cctgatatc 1680
accctgatgc aggaagacca gaaccagcag gagaagaagc agcagagct gctggagctg 1740
gacaagtggc caagcctgtg gaacctgttc gacatcagca agtggtgtg gtacatacag 1800
atcttctcac tgatcgtggg cggcctgggt ggcctgcgca tegtgttacc cgtgtgagc 1860
atcgtgaacc gcgtgcgcca gggtcatagc cccctgagct tccagaccg cttccccgcg 1920
cccccgggcc cgaccgcccc cgaggcgatc gaggagagcg cggcgagcg cgaccgcac 1980
cgcgacgacc ccttggtgca cggcctgtgt gccctgatct gggcgagacc gcgcagcgt 2040
tgccctgtca cgtaccacgc cctgcgcgac ctgatcctga tgcgcggcg catcgtgagg 2100
ctgctggggc gcccgcggtg ggaggccctg aagtactggg gcaactcgtc gcagtactgc 2160

```

```

atccaggagc tgaagaacag cgccgtgagc ctgttcgagc ccacccgcat cgccgtggcc 2220
aggggcaacc accgcatcat cgaggtggcc cagcgcacgc gcgcgcctt ctgcacate 2280
ccccgcgcga tccgcacagg ctctgagcgc gccctgetgt aactcgagcg tgct 2334

```

```

<210> 8
<211> 2316
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: Val120-Thr202

```

```

<400> 8
gaattcgcca ccattggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
cgagttcttc ttctgcccag cgccgtggag aagctgtggg tgaccgtgta ctaccgctgc 120
cccggttgga aggaaggccac caccaccctg ttctgcccga gcgacgcaa ggctctcgac 180
accgaggtgc acaacgtgtg ggccaccac cgcctgcgtg ccaccgaccc caaccgccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catgtgggag 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtggggcgc 360
gccaccacgg cctgcccacaa ggtgagcttc gagcccatcc ccactcccta ctgcgcccc 420
gcgggcttgc ccactcctgaa gtgcaacgac aagaagtcca accggcagcg ccctcgacc 480
aacgtgagca cgtgcagtg caccacggc atccgcccc tggtgagcac ccagctgctg 540
ctgaagcgga gccctggcgga ggaggcgctg gtgatccgca gcgagaactt caccgacaa 600
gccaaagcca tcactgtgca gctgaaggag agcgtggaga tcaactgcac ccgcccacac 660
aacaacaccc gcaagagcat caccatcgcc ccgggccgcg ccttctacgc caccggcgac 720
atcatcgggc acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagttcgcca acaagacctt cgtgttcaag 840
cagagcagcg gcggcgaccc cgagatcgtg atgcacagct tcaactgcgg cggcgagttc 900
ttctactgca acagacacca gctgttcaac agcacctgga acaacacctt cgccccacac 960
aacaccacgc gcaccatcac cctgcctgc cgcataaagc agatcatcaa ccgtggcgag 1020
gaggtgggga aggcactgta ccccccccc atccgcggcc agatccgctg cagcagcaac 1080
atccacggcc tgcgtctgac ccgcgacggc ggcaaggaga tcagcaaac caccagatc 1140
ttccgcccc gcggcgcgga catgcgcgac aactggcgca gcgagctgta caagtacaag 1200
gtgggtgaaga tcgagccctt gggcgtggcc ccaccaaagg ccaagcgccg cgtgtgcaag 1260
cgcgagaagc gcgcgctgac cctggggccc atgttctctg gcttctctgg gcgcgcggc 1320
agcacctagg ggcggcgag cctgacccct acgctgcagg ccgcccagct cgtgagcgcc 1380
atcgtgcagc agcagaacaa cctgctgcgc gccatcgagg ccacagcga cctgctgcag 1440
ctgacgtgtt ggggcatcaa gcagctgcag gccccgctgc tggccgtgga gcgctacctg 1500
aaggaccagc agctgctggg catctggggc tgacgcgga agctgatctg caccaccgcc 1560
gtgcccctga acgcccagct gagcaacaag agcctggacc agatctggaa caactagacc 1620
tggtggagtg tggagcgcca gatcgacaac tacaccaacc tgactacac ccctgacgag 1680
gagagcaga accagcagga gaagaacgag caggagctgc tggagctgga caagtgggcc 1740
agcctctgga actggttcga catcagcaag tggctgtggt acatcaagat cttcatcatg 1800
atcgtggggc gctggtgggg cctgcgcata gtgttcacgc tgctgagcat cgtgaaccgc 1860
gtgcggcagg gctacagccc cctgagcttc cagacccgct tcgccgcccc ccggcgcccc 1920
gacgcggccc agggcatcga ggaggaggcg ggcgagcgcg accgcgacgc cagcagcccc 1980
ctggtgcagc gctgctggc cctgatctgg gacgacctgc cgacctgtg cctgttcaag 2040
taccacggcc tcgcgcacct gatcctgat cgcgcggcca tcgtggagct gctggggcgc 2100
cgcgctggg aggcctgtaa gtactggggc aaactgctgc agtactggat ccaggagctg 2160
aagaacagcg cgtgagcctt gctcagcgcc atcgccatcg cgtggcgga gggcaccgac 2220
ggcatcatcg aggtggccca gcgcacggc cgcccttccc tgcacatccc ccgcgcgac 2280
cgccagggtc tcgagcgcg cctgctgtaa ctgagag 2316

```

```

<210> 9
<211> 2541
<212> DNA
<213> Artificial Sequence

```

```

<220>

```

## &lt;223&gt; Description of Artificial Sequence: Trp427-Gly431

&lt;400&gt; 9

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagctcttcg ttctgcccgac cgcctgtggag aagctgtggg tgacctgtga ctacggcgctg 120
cccgctgtgga aggaggccac caccacctctg ttctgcgcca gcgacgccaa ggctctacgac 180
accgaggtgc acaacgtgtg ggcaccccaac gctctgctgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaaatgtt ggaagaacaa catgtgtggag 300
cagatgcacg aggcacatcat cagcctgtgg gaccagagcg tgaagccctg cgtgaaagctg 360
acccccctgt ggtgacacct gcactgcacc aacctgaaga accgacacaa cccaaagagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaagt gcagettcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gcctctgtct acaagctgga cgtggtgccc 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagctg gatcaccacg 600
gctgcgccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccgcttc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gccctgcac caactgagc 720
accgtgcagt gcaaccacgg catccgcccc gtgtgtgagc cccagctgct gctgaaacggc 780
agcctggcgc aggaggcggt ggtgatccgc agcgagaact tcaccgacaa caccggcacc 840
atcatctgctc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccgccgcg gctcttaacg ccaccggcga catcatcggc 960
gacatccgcc agggccaactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcgtgacca agctgcagcg ccagttcggc aacaagacca tctgtgtcaa gcagagcagc 1080
ggcgggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt ctctactgc 1140
aacagcacc cagctgtcaa cagcaccctg aacaacacca tgggccccaa caacaccac 1200
gtgacataca ccttgcctcg ccgcatcaag ccgctgggg accgaaggcc 1260
atgtacgccc cccccatcgc cggccagatc cgtcgagca gcaacatcac cggctcgtc 1320
ctgaccctgc acggcgccaa ggagatcagc aacaccacgy agatcttcgy ccccggcgcc 1380
ggcgacatgc ggcgaactg gcgcagcgag ctgtacaagt acaagtggt gaagatcgag 1440
ccccctggcg tggccccac caaggccaa gcccgcgtgg tgacgcgca gaagcgcgc 1500
gtgacctggc ggcctatgtt cctgggcttc ctggggccgy cggcgagcac catggggcc 1560
cgacgtgca cctgacacct gcaaggccgc cagctgctga gcggcatcgt cgacgacgag 1620
aacaacctgc tgcgcacctc cagggccacg cagcaccctg tcagctgac cgtgtggggc 1680
atcaagcagc tgcaggcccg cgtgctggcc gtggagcgtg acctgaagga ccagcagctg 1740
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgctgtgc ctggaacgac 1800
agctggagca acaagagcct gaaccagatc tggaaacaa tgacctggat gaggctggag 1860
cgcgagatcg acaactacac caacctgac tacacctgta tcgaggagag ccagaaccag 1920
caggagaaga acgagcagga gctgctggag ctggacaagt gggccagcct gtgaaactg 1980
ttcgacatca gcaagtggct gtgtacatc aagatcttca tcatgatcgt gggcggtcgt 2040
gtgggctcgc gcatcgtgtt caccgtgctg agcatcgtga acccggtgcg ccagggtcac 2100
agccccctga gcttccagac ccgcttcccc gcccccgcg gccccgacgy ccccagggc 2160
atcgaggagg agggcgcgga gcgcgacggc gaccgcagca gccccctggt gcaacggctc 2220
ctggccctga tctggagca cctgcgcagc ctgtgctgtg tcaactacca ccgctcgcg 2280
gacctgaccc tgatgcggcg ccgcatcgtg gagctgctgg gcgcgcgcy ctggaggacc 2340
ctgaagtact ggggcaacct gctgcagtac tggatccagg agctgaagaa cagcgccgtg 2400
agcctgtctg acgcatcgc catcgccgtg gccaggggca ccgacccgat catcgaggt 2460
gccagcgca tcggccgcgc ctctctgcac atcccccgcc gcacccgcca gggcttcgag 2520
cgccgctcgt tgtaactcga g 2541

```

&lt;210&gt; 10

&lt;211&gt; 2541

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Arg426-Gly431

&lt;400&gt; 10

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagctcttcg ttctgcccgac cgcctgtggag aagctgtggg tgacctgtga ctacggcgctg 120
cccgctgtgga aggaggccac caccacctctg ttctgcgcca gcgacgccaa ggctctacgac 180

```

```

accgaggtgc acaacgtgtg ggcacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaaact gaccgagaac ttcaacatgt ggaagaacaa catggtgtgg 300
cagatgcacg aggacatcat cagcctgtgg gaccagagcc tgaagccctg ctgtaagctg 360
acccccctgt gcgtgaccct gcactgcacc aacctgaaga acgcccacaa caccagaagc 420
agcaactgga agagatgga ccgcggcgag atcaagaact gcaagctcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gcctctgtct acaagctgga cgtgtgtccc 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcacccag 600
ccctgcacca aggtgagctt cgagcccatc cccatccact actgcgcccc cgccgggttc 660
gccatcctga agtgcacaga caagaagttc aacggcagcg gccctgcac caagctgagc 720
accgtgcagt gcacccacgg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agcctggccg aggagggcgt ggtgatccg agcagagact tcaccgacaa cgccaagacc 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca ccgcgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccgggcgc gcctttacg ccaccggcga catcatcgcc 960
gacatccgca agcccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcgtgacca agctgcaggc ccagttcgcc aacaagacca tcgtgttcaa gcagagcagc 1080
ggcggcgacc agctgatcgt gatgcacagc ttcaactgcg gcggcgagtt ttctactgct 1140
aacagacacc agctgttcaa cagcacctgg aacaacacca tcggcccaaa caacaccac 1200
ggcaccatca cctgcgctcg ccgcatcaag catgatcatc accgcgcgcg cggaagggcc 1260
atgtacgccc cccccatccg cggccagatc cgctgcagca gcaacatcac cgcctcgtgt 1320
ctgaccgccg acggcgggcaa ggagatcagc aacacacccg agatcttccg ccccgggcgc 1380
ggcgacatgc gcgacaactg gcgcagcgag ctgtacaagt acaaggtgtt gaagatcgag 1440
ccccgtggcg tggcccccac caaggccaa gcccgcgtgt tgcagcgca gaagcgcgcc 1500
gtgaccctgg gcgcattgtt cctgggcttc ctggggcgcc cggcgagcac catggggcgc 1560
cgacatcgta cctgacccgt gcaggccgcg cagctgctga gcggcatcgt gcagagcagc 1620
aacaacctgc tgcgcgccat cagggcccgag cagcacctgc tgcagctgac cgtgtggggc 1680
atcaagcagc ctgagggccc cgtgctggcc gtggaagcgt acctgaagga ccagcagctg 1740
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgcgctgcc ctggaacgcc 1800
agctggagca acaagagcct ggacagatc tgggaacaaa tgaactggat ggaagtggag 1860
cgcgagatgc acaactacac caacctgatc tacacccatg tcagggagag ccagaaccag 1920
caggagaaga caggcagga gctgctggag ctggacaagt gggccagcct ctggaactgc 1980
ttcgacatca gcaagtgcgt gtggtacatc aagatcttca tcagtatcgt gggcggcctg 2040
gtggccctgc gcatcgtgtt caccgtgctg agcatcgtga accggtgcg ccagggtac 2100
agccccctga gtttccagac ccgcttcccc gcccccgcgc gcccgacgc ccccgaggcc 2160
atcgaggagg agggcgcgca gcgcgacgcg gacgcgagca gccccctggt gcacggcctg 2220
ctggccctga tctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgctgcgc 2280
gacatcctc tgatcgccgc ccgcatcgtg gagctgctgt gccgcgcgcg ctgggagagc 2340
ctgaagctac ggggcaacct gctgcagtac tggatccagg agctgaagaa catcgccgtg 2400
agcctgttgc accgcatcgc catcgccgtg gccagaggga ccgacgcgat catcgaggtg 2460
gcccgagcga tccggcgcg cttcctgcac atccccgcgc gcatccgca gggcttcgag 2520
cgcgccctgc tgaactcga g

```

&lt;210&gt; 11

&lt;211&gt; 2541

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Arg426-Gly431E

&lt;400&gt; 11

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gttgtgctgt gctgtgtgga 60
cgagtcctgc ttctgcccag cgcgctggag aagctgtggg tgaccgtgta ctacggcgtg 120
ccgctgtgga aggagggcac caccaccctg ttctgcgcca gcgacgcca ggctacagca 180
accgaggtgc acaacgtgtg ggccacccac gcctgcgtgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacct gaccgagac ttcaacatgt ggaagaacaa catggtggag 300
cagatgagca aggacatcat cagcctgtgg gaccagagcc tgaagccctg ctgtaagctg 360
acccccctgt gcgtgaccct gcactgcacc aacctgaaga acgcccacaa caccagaagc 420
agcaactgga aggagatgga ccgcggcgag atcaagaact gcaagcttca ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gcctctgtct acaagctgga cgtgtgtccc 540

```

```

atcgacacag acaacaccag ctacaagctg atcaactgca acaccagcgt gatcacccag 600
gccctgccca aggtgagctt cgagcccatc cccatccact actgcgcccc cgcgggcttc 660
gccatctcga agtgcaacga caagaagttc aacggcgagc gccctctcac caactgtgagc 720
accgtgcagt gaccccaacg gatccgcccc gtggtgagca cccagctgct gctgaaaggc 780
agcctggcgc agggagggcgt ggtgatccgc agcgagaaat tcaccgacaa cgccaagacc 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccgccgcg gccctctacg ccaccggcga catcatcggc 960
gacatccgcc agggccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcgtgacca agctgcaggc ccagtctggc aacaagacca tcgtgttcaa gcgagcagc 1080
ggcgggcagc ccgagatcgt gatgcacagc ttcaactgcg gcggtgagtt cttctactgc 1140
aacagacacc agctgttcaa cagcacctgg aacaacacca tcggccccaa caacaccaac 1200
ggcaccatca cctgcgccct ccgcatcaag catgatcatca acccgggcag cggcaaggcc 1260
atgtacgccc ccccacatcg cggccagatc cgtgcagca gcaacatcac cggcctgctg 1320
gtgacccgcg acggcggaag gagatcagc aacaccaccg agatcttccg ccccgggcgc 1380
cgcgacatgc gcgacaactg gcgcagcgag ctgtacaagt acaaggtggt tcaagatcag 1440
ccctggggcg tggcccccac caaggccaaag cgcgcgtggt tgcagcgaga gaagcgcgcc 1500
gtgacctcgt cggcatgttt cctgggcttc ctgggcgccg ccggcagcac catggggcgc 1560
cgacagctga ccttgacctg gcaggccccc cagctgctga gcggcatcgt gcagcagcag 1620
aacaacctgc tgcgcgccat cgaggccccc cagcacctgc tgcagctgac cgtgtggggc 1680
atcaagcagc tgcaggcccg cgtgtgggct gtggagcgct acctgaagga ccagcagctg 1740
ctgggcatct ggggtgcgag cggcaagctg atctgcacca ccgccgtgcc ctggaagcgc 1800
agctggagca acaagagcct ggaccagatc tggaacaaca tgacctggat agagtgggag 1860
cgcgagatcg acaactacac caacctgatc tacacctga tcgaggagag ccagaaaccg 1920
caggagaaga agggcagcga gctgctggag ctggacaagt gggccagcct gtggaaactg 1980
tctgcacatca cgaagtgcct gtggtacatc aagatcttca tcatgatcgt gggcgggcct 2040
gtggggcctg cactcgtggt caccgtgctg agcatcgtga accgctgctg caggcgctac 2100
agcccccctg gcttccagac ccgcttcccc gcccccgcg ccccgccagc ccccgagggc 2160
ctcggaggag agggcgcgga gcgcgacgcg gacgcgcaga gccctcgtgt tggggagcgt 2220
ctggccctga tctgggaaga cctgcgcagc ctgtgcctgt tcagctacca ccgctcgcg 2280
gacctgatcc tgatcgccgc ccgcatcgtg gagctgctgg gccgcgcggy ctggggagcg 2340
ctgaagtact ggggcaacct gctgcagtac ttgattccag agctgaagaa cagcgccgctg 2400
agcctgttgc acggcatcgc catcgccgtg gccgaggga ccgaccgcac catcgaggtg 2460
gcccgaccga tcggccgcgc ctctctgcac atcccccgcc gatccgcca gggcctcgag 2520
cgcgcctcgc tgtaactcga g

```

2541

&lt;210&gt; 12

&lt;211&gt; 2541

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Arg426-Lys432

&lt;400&gt; 12

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
cgagctcttg ttctgccagc cgcgctggag aagctgtggg tgacctgtga ctacggcgtg 120
ccgctgtgga aggaggccac caccacctgt ttctgcgcca gcgacgacca ggccctacgac 180
accgaggtgc acaactcgtg gccaccacac gcctcgtcgt ccaccgaccc caacccccag 240
gagatcgtgc tggagaacct gaccgagaac ttcaacctgt ggaagaacaa caatgtggag 300
cagatgcacg aggacatcat cagcctctggg gaccagagcc tgaagccctg cgtgaagcgt 360
acccccctgt gcgtgacctg gcactgcacc aacctgaaga acgcacacca caccaaagac 420
agcaactcga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct caaagctgga cgtgtgccc 540
atcgacaacg acaaacaccg ctacaagctg atcaactgca acaccagcgt gatcacccag 600
gcttcgcccc aggtgagctt cgagcccatc cccatccact actgcgcccc cgcgggcttc 660
gccatcctga agtgcaacga caagaagttc aacggcgagc gccctctcac caactgtgagc 720
accgtgcagt gaccccaacg gatccgcccc gtggtgagca cccagctgct gctgaaaggc 780
agcctcgagg agggagggcgt ggtgatccgc agcgagaaat tcaccgacaa cgccaagacc 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgccccaa caacaacacc 900

```

```

gcacaagagca tcaaccatcgg ccccgccggc gcttctacg ccaccggcga catcatcggc 960
gacatccgcc aggeccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcgtgaccca agctgcaggc ccagttcggc aacaagacca tctgtgtcaa gcagagcagc 1080
ggcgggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcgggcaggt ttcttactcg 1140
aacagacccc agctgttcaa cagcacctgg aacaacacca tcggcccccac caacacaaac 1200
ggcaccatca cctgtccctg ccgcatcaag cagatcatca accgcggcgg caacaaggcc 1260
atgtacggccc ccccacatcg cggccagatc cgctgcagca gcaacatcac cggcctgctg 1320
ctgaccccgg acggcggcga ggaagatcag aacaccaccc agatcttccg ccccgcgccg 1380
ggcgacatgc gcgacaactg gcgcagcgag ctgtacaagt acaagtggtt gaagatcgag 1440
ccctcggggc tggcccccac caaggccaaag cgcgcgtgtg tgacgcgcga gaagcgcgcc 1500
gtgacctctg cctgactgtt cctgggcttc ctgggcgcgc ccgggcgcgc 1560
cgacagcctga cctgacgctg gcaggcccg cagctcgtga gcggcatcgt gcagcagcag 1620
aacaaactgc tgcgcgccat cgaggcccg cagcacctgc tgcagctgac cgtgtggggc 1680
ctcaagcagc tgcaggcccg cgtgctggcg gtggagcgtt acctgaagga ccagcagctg 1740
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccgcctgccc ctggaacgcc 1800
agcttgagca acaagagcct ggaccagatc tggacaacca tgacctggat ggaagtggag 1860
cgcgagatgc acaactacac caacctgac tacacctga tcgaggagag ccagaacacc 1920
caggagaaga acgagcagga gctgctggag ctggacaagt gggccagcct gtggaactcg 1980
ttcgacatca ccaagtggct gtggtacatc aagatcttca tcatgatcgt cggcgccctg 2040
gtgggctctg gcatcgtgtt caccgtgctg agcatcgtga accgctgctg ccagggtctc 2100
agcccccctg gcttccagac ccgcttcccc gccccccggc gcccgcagcc ccccgaggcg 2160
atcgaggagg agggccggca gcgcgacccg gacgcgagca gcccctgggt gacccggctg 2220
ctggccctga tctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgcctgcgc 2280
gaactcatcc tctggccccc cgcgacgtg gagctcgtgg gcgcgcgcgc ctggagcgcc 2340
ctgaagtact ggggcaacct gctgcagtac tggatccagg agctgaagaa cagcgcgctg 2400
agcctgttgc ccaactcgc catcgcctg gcgcagggca ccgacgcgat catcgaggtg 2460
gccacgcgca tccggccgcg ctctcctgac atcccccgcc gcatccgcga gggcttcgag 2520
cgcgcctctg tgtaactcga g

```

<210> 13  
 <211> 2535  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Asn425-Lys432

```

<400> 13
gaattcgcca ccatgggatc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
cgagttcttg ttctgcccag cgcctgtggg aagctcgtgg tgacctgta ctacggcgtg 120
ccggtgttga agaggccac caccacctgt ttctgcgcca gcgacgcgca ggcctactag 180
accgaggttg acaacgtgtg ggccaccacc gctcgtgctg ccaccgaccc caacccccag 240
gagatcgtgc tggagaacct gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcagc aggacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acgccctctg gctgacctct gcactgcacc aacctgaaga acgcacacca accaagaagc 420
aggcaactga aggagatgga ccgcggcgag atcaagaact gcagctcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtggtgcc 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagcgt gatcacccg 600
gcttgcgcca aggtgagctt cgagcccact cccatccact actgcgcccc cgtggctctc 660
gccatcctga agtgcaacga caagaagttc aacggcagcg gccctcgcac caactgagc 720
accgtgcagt gcaaccacgg catccgcccc gtggtgagca ccagctgct gctgaacgct 780
agcctggcgg aggaggcggt ggtgatccgc agcgagaact tcaccgacaa ccgcaagacc 840
atcaactgtg agctgaagga gagcgtggag atcaactgca cccgcccacaa caacaacacc 900
cgcaagagca tcaccatcgg cccggccggc gcttctacg ccaccggcga catcatcggc 960
gacatccgcc agggccactg caacatcagc ggcgagaagt ggaacaacac cctgaagcag 1020
atcgtgacca agctgcaggc ccagttcggc aacaagacca tctgttcaa gcagagcagc 1080
ggcggtgcac ccgagatcgt gatgcacagc ttcaactcgc gcggcgagtt ctcttactcg 1140
aacagacccc agctgttcaa cagcacctgg aacaacacca tcggcccccac caacacaaac 1200
ggcaccatca cctgcctctg ccgcatcaag cagatcatca acgcccccaa ggccatgtac 1260

```

```

gccccccca tccgcggcca gatccgtctgc agcagcaaca tcaccggcct gctgtgtgacc 1320
cgcgacggcg gcaaggagat cagcaacacc accgagatct tccgcccccg cggcgggcgac 1380
atgcggcgaca actggcgagc cagcgtgtac aagtaacaagg tggtagaagt ctacggccctg 1440
ggcgtggccg ccaccaaggc caagcgccgc gtggtgcagc gcgagaagcg cgcgctgacc 1500
ctggggcgcca tgttccctggg ctccctgggc gcccgccgga gccacatggg cggccgcgacg 1560
ctgaccctga cgtgcagcgc ccgccagctg ctgagcgcca tcgtgcagca gcagaaacac 1620
ctgtcgcgcg ccacagaggg ccagcagcac ctgctgcagc tgaccgtgtg gggcgtcaag 1680
cagctcgagg ccgcgctgtc ggcctggagg cgtacctga agaccagca gctgctgggg 1740
atctggggcgt ccagcgcgaa gctgatctgc accaccggcg tgccctggaa gctcagctgg 1800
agcaacaaga gctcgagcca gatctggaac aacatgacct ggaatggagt ggaagcgag 1860
atcgacaact accacaacct gatctacacc ctgctcgagg agaccagaa ccagcaggag 1920
aagaacgagc aggaagctgt ggagctggac aagtgggcca gcctgtggaa ctggttcgac 1980
atcagcaagt ggcttggtta catcaagatc ttcatcatga tcgtggcgcg cctggtggcg 2040
ctgcgcctgc tgttccaccg gctgagcacc gtgaaccgcg tgcgccaggg ctacagcccc 2100
ctgagcttcc agaccgcttt ccccgccccc cgcggccccc accgccccga gggcactcgag 2160
gaggagggcg gcgagcgcca ccgcgaccgc agcagccccc tggtagcagg cctgtggcg 2220
ctgacttggg ccagcctgcg cagcctgtgc ctgttcagct accaccgctc cgcgaagctg 2280
atcctgatgc ccgcccgcat cgtgagagct ctgggcccgc cggcgctggg ggcctctgag 2340
tactggggca acctgtgca gtactggatc caggagctga agaaccagcg cgtgagctgc 2400
ttcgagcgca tgcgcatgc cgtggcgagc ggcaccgacc gcacatcga ggtggccag 2460
cgcatcgccc gcgctctctt gcacatcccc cgcgcctacc gccagggctt cgagcgcgcc 2520
ctgctgtaac tcgag                                     2535

```

<210> 14  
 <211> 2529  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Ile424-Ala433

```

<400> 14
gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagctcttcg ttctgcccag cgcgctggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgctgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcca ggcctacgac 180
accgaggtgc acaacgtgtg ggccaccac gccctcgtgc cccaccgacc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catgtgtggag 300
cagatgcagc aggaatcatc cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
acccccctgt cgttgacctc gcactgcacc aacctgaaga acgcccacaa caccagagc 420
agcaacttga aggagatgga ccgcggcgag atcaagaact gcagcttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gcctgtttct acaagctgga cgtggtgccc 540
agtgacaacg acaaccacag ctacaagctg atcaactgca acaccagctg gatcacccag 600
gcttgcccga aggtgagctt cgagcccatc cccatccact actgcgcccc cgcgggcttc 660
gccatcctga agtgacaag caagaagttc aacggcagcg gccctgcac caacgtgagc 720
accgtgcagt gcaccacagg catccgcccc gtggtgagca cccagctgct gctgaacggc 780
agcctggcgg aggaggcgtt ggtgatccgc agcagaaact tcaccgcaaa ccgccaagac 840
atcatcgtgc cgtgaagga gagcgtggag atcaactgca cccgcccac caacaacac 900
cgcaagagca tcaccactcg ccccgccgcg gcccttctacg ccaccggcga catcatcgcg 960
gacatccgcc agggccactg caacatcagc ggcgagaagt ggaacaacac cctggaagcag 1020
atcgtgacca agctgcagcg ccagttccgc aacaagacca tcgtgttcaa gcgagcagc 1080
ggcgggcgacc cggagatcgt gatgcacagc ttcaactgcg gcggcgagtt ctcttactgc 1140
aacagcacc cagctgttoaa cagcaactcg aacaacacca tcggccccaa caacaccac 1200
ggcaccatca cccctgcccgt ccgcacaaag cagatcatcg gcggcgccat gtagcgcccc 1260
gccatccgcg ccagatcccg ctgcagcagc aacatcacgc gctcgtctgt gaccgcgac 1320
ggcgggcaag agatcagcaa caccaccagc atcttcggcc cggggcgggc cgacatcg 1380
gacaactggc gcagcagact gtacaagtac aagtggtgga agatcgagcc cctgggtgtg 1440
gccccacaca agggcaagcg ccgctggtg cagcgcgaga agcgcgcgtg gacccctggc 1500
gccatgttcc tggcttctct gggcgccgcg ggcagacca tggcgcccg cagcctgacc 1560
ctgacgtgac agggccgcca gctgctgagc ggcactcgtg agcagcagaa caacctgctg 1620

```



```

cgccgcatcg agggccagca gcaactgctg cagctgaccg tgtggggcat caagcagctg 1680
cagggcccgcg tctgtggcgt ggagcgctac ctgaaggacc agcagctgct gggcactctg 1740
ggctgcagcg gcaagctgat ctgacaccac cgcgtgcctt ggaacgctcg ctggagcaac 1800
aagagcctcg accagatctg gaacaacatg acctggatgg agtgggagcg cgagatcgac 1860
aactacacca acctgatcta caccctgata gaggagagcc agaaccagca ggagaagaac 1920
gagcagagagc tctgtgagct ggacaagtgg gccagcctgt ggaactgggt cgacatcagc 1980
aagtggcgtgt ggtacatcaa gatcttcata atgatcgtgg gcggcctggt gggcctgcgc 2040
atcgtgttca ccgtgctgag catcgtgaac cgcgtgcgcc agggctacag cccctgagc 2100
ttccagaccc gtttccccgc ccccgccggc cccgagcgct cgaggagcgt 2160
ggcgggcgagc ggcagccgca ccgcagcagc cccctggctg acggcctgct ggcctgatc 2220
tgggagcacc tgcgagcct gtgcctgttc agctaccacc gctgcgcga cctgatcctg 2280
atcgccgccc gcatcgtgga gctgctggcg cgcgcggctt gggagggcct gaagtaactg 2340
ggcaacctgc tgcagtactg gatccaggag ctgaagaaca cgccgtgag cctgttctgc 2400
gccatcgcca tgcgcgtggc cgagggcacc gaccgcatca tcgaggtggc ccagcgcatc 2460
ggccgctgct tcctgcacat ccccgccgcg atccgcagcg gcttcgagcg ccgctctgct 2520
taactcgag 2529

```

&lt;210&gt; 15

&lt;211&gt; 2523

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Ile423-Met434

&lt;400&gt; 15

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
cagctcttctg ttctgcccaag cgcctgggag aagctgtggg tgaccgtgta ctacggcggtg 120
cccggtgtgga aggagggcacc caccaccctg ttctgcgcca ggcagcgcaa ggctcaagac 180
accgaggtgcg acaacgtgtgt ggccaccacc gctcgtgtgc ccaccgaccc caacccccagc 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catgtgtggag 300
cagatgcacg aggcacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaaactg 360
acccccctgt gctgagccct gcactgcacc aacctgaaga acgcccacaa caccaaagag 420
agcaactgga agagatgga ccgcggcgag atcaagaact cgacttcaa ggtgaccacc 480
agcatccgca acaagatgca gaaggagtac gccctgttct acaagctgga cgtgtgtccc 540
atcgacaaag acaacaccag ctacaagctg atcaactgca acaccagcgt gatcaccacc 600
gctgccccca aggtgagctt cgagcccacc ccatccaact actgcgcccc cgcgcggctt 660
gccatcctga agtgcaacga caagaagtct aaccgagcgg gccctgcac caacgtgagc 720
accgtgcagt gcaccacacg catccgcccc gtgtgtgagca cccagctgct gctgaacgag 780
agcctggcgg aggagggcgt ggtgatccgc agcgagaact tcaccgacaa cgccaagacc 840
atcatcgtgc agctgaagga gagcgtggag atcaactgca cccgcccacaa caacaacacc 900
cgcgaagaga tcaccatcgg ccccgccgcg gccttctatc ccaccgcgga catcatcgcg 960
gacatccgcc agggccactg caacatcagc ggcgagaagt ggaacaacac ctgatcagcg 1020
atcgtgacca agctgcagcg ccagttcggc aacaagacca tcgtgttcaa gcagagcagc 1080
ggcgcgagcc ccgagatcgt gatgcacagc ttcaactgcy gcggcgagtt ctctactgc 1140
aacagcaccc agctgttcaa cagcacctgg aacaacacca tcggcccacaa caacaccac 1200
ggaccatcaa cctgcacctg ccgcatcaag cagatcgggc gcatgtacgc cccccccatc 1260
cgcggccaga tcctcgtcag cagcaaacat accgcctgtc tgcgtgacccg cgacggcgcg 1320
aaggagata gcaaacaccac cgagatcttc cgcgccggcg gcggcgacat gcgcgcaaa 1380
tggcgacagc agctgtacaa gtacaaggtg gtgaagatcg agccctggg cgtgcgcccc 1440
accaagcgca agcgcgcgtt ggtgcagcgc gagaagcgcg cctgaccctt ggcgcgcatg 1500
ttctgggct tcttggggcg cgcggcgagc acctggggcg cccgcagcct gaccctgacc 1560
gtcgagggcc gccagctgct gagcggcctc gtgcagcagc agaacaacct gctgcgccc 1620
atcgagggcc agcagacctt gctgcagctg acogtgtggg gcatacaaga cgtgcagggc 1680
cgcgtctgct ccgtggagcg ctacctgaag gaccagcagc tgcctgggcat ctggggctgc 1740
agcggccaaag tgatctgcac caccggcgtg ccttggaaag ccagctggag caacaagag 1800
ctggaacaga tctggaacaa catgacctgt atggagtggg agcgcgagat cgacaacctc 1860
accaacctga tctacacctt gatcgaggag agccagaacc agcaggagaa gancgagcg 1920
gagctcgtcg agctggacaa gtggggcagc ctgtggaaat ggttcgacat cagcaagtgg 1980

```

```

ctgtggtgaca tcaagatctt catcatgac gtggggcgcc tgggtggcct ggcacatcg 2040
ttaccgctgc tgagcatcgt gaaccgcgtg gccacgggct acagcccctg gacttccag 2100
accgcttccc ccgccccccg cggccccgag cggccccgag gcatcgagga ggagggcgcc 2160
gagcgcgacc ggcagccgag cagccccctg gtgcacggcc tgcctggcct gatctgggac 2220
gacctgcgca gctgtgctct gttcagctac caccgctcgc gcgacctgat cctgatcgcc 2280
cccgctcgtg tggagctgct gggcccgccg ggctgggagg ccttgaagta ctggggcaac 2340
ctgctgcagt actggatcca ggagctgaag aacagcgccg tgagcctgtt cgagcccat 2400
gccatcgccg tggccgaggg caccgacccg atcatcgagg tggccacagc catcgccgcg 2460
gccttctctg acatcccccg ccgcatcccg cagggtctcg agcgcgccct gctgtaactc 2520
gag 2523

```

&lt;210&gt; 16

&lt;211&gt; 2517

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Gln422-Tyr435

&lt;400&gt; 16

```

gaattcgcca ccatggatgc aatgaagaga gggtctctgt gtgtgtctgt gctgtgtgga 60
cgagtctctc ttccgccag cccgtggag aagctgtggg tgacctgtga ctacggcgtg 120
cccgtgtgga aggagggcac caccacctgt ttctgtgcca gcgacgccaa ggctacgcac 180
accgaggtgc aacacgtgtg ggccaccac cctctgctgc ccacgcaccc caacccccac 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catgtgtggag 300
cagatgcacg aggaacatcat cagcctgtgg gaccagagcc tgaagccctg cgtggaagctg 360
ccccctctgt cgtgtgacct gcactgcacc aacctgaaga acgccaccaa caccagagc 420
agcaactgga aggagatgga ccgcccgcag atcaagaact cgaacttcaa ggtgacacac 480
agcatccgca acaagatgca gaaggagatc gccctgttct acaagctgga cgtggtgccc 540
atcgacaacg acaacaccag ctacaagctg atcaactgca acaccagctg gatcaccacg 600
gcttccccca aggtgagctt cgagcccatc cccatccact actgcgcccc ccgctggctt 660
gccatcctga agtgcaacga caagaagttc aacggcgagc gccctgcac caactgtgagc 720
accgtgcagt gcacccaacg catccgcccc gtggtgagca cccagctgct gctgaacgcg 780
agcctggccc aggagggcgt ggtgatccg agcgagaact taccgcgcaa gcccaagacc 840
atcatcgtgc agctgaaaga gagcgtggag atcaactgca cccgccccaa caacaacacc 900
cgcaagagca tcaccatcgg ccccgcccg gccttctac ccacggcgga catcatcgcc 960
gacatccgac agggccactg caacatcagc ggcgagaagt ggaacaacac ctctgaagcag 1020
atcgtgacca agctgcaggg ccagttccgg aacaagacca tctgtttcaa gcagagcagc 1080
ggcgggcgacc ccgagatcgt gatgcacagc ttcaactgcg gcggcgagtt ctctactgc 1140
aacacacacc agctgttcaa cagcaccctg aacaacacca tcggccccaa caacccaacc 1200
ggcaccatca ccttgcctcg ccgatcaag cagggcggtc acgccccccc catccgcgcg 1260
cagatccgct gcagcagcaa catcacccg ctgctcgtga cccgcagcgg ccgcaaggag 1320
atcagacaac ccacgagat ctctcgcccc ggccggcgcg acatcgcgca caacttgcgc 1380
agcgagctgt acaagatcaa ggtggtgaa atcgagcccc tggcgctggc ccccaaccaag 1440
gccaaagccc gctgtgtgca gcgcgagaag ccgcccgtga cctggggcgc catgttccctg 1500
ggcttctctg gcgcgcggcg cagcaccatg ggccggcgca gctgacctg gacgtgcag 1560
gcgcgcacgc tgctgagcgg catcgtgcag cagcagaaac acctgtgcg ccctatcgag 1620
gcccgacgac acctgctgca gctgacctg tggggcatca agcagctgca ggcccggtg 1680
ctggcgctgg agcgctacct gaaggacacg cagctgctgg gactgtggg gactgcggc 1740
aagctgatct gcaccaccgc cgtgccctgg aacgcagct ggagacaaca gagctctggac 1800
cagatctgga acaacatgac ctggatggag tggagagcgc agatgcacaa ctacaccaac 1860
ctgatctaca cctgatgaca ggagagccag aaccacgag agaagaacga gcaggagctg 1920
ctggagctgg acaagtgggc cagcctgtgg aactggctcg acatcagaa ctgggtgtgg 1980
tacatcaaga tcttcaatcat gatcgtgggc ggcctgggtg gctgcgcat cgtgttacc 2040
gtgctgagca tctgtaaacg cgtgcgccag ggctacagcc ccttgagctt ccagaccgcg 2100
ttccccgcgc ccccggcccc cgaccggccc gaggggcatc aggagggagg ccgagagcgc 2160
gaccgcgacc gcagcagccc cctgggtgac ggctcgtgag cctgatctg ggaagacctg 2220
cgagacctgt gctgttcaag ctaccaccgc ctgcgcgacc tgactcctgat cgcgcgcgcg 2280
atcgtggagc tgcgtgggcg ccgcggtctg gaggccctga agtactgggg caacctgctg 2340

```

```
<210> 17
<211> 2517
<212> DNA
<213> Artificial Sequence
```

gaattc	17	ccatggatgc	aatgaagaga	gggctctgct	gtgtgctgct	gctgtgtgga	60
gcagctcttc		tttcgccccc	cgccgtggag	angctgtgtg	tgacgcgtga	cttcggcgctg	120
ccccgttggga		acgaagccac	caaccacctt	ttctgcgcga	gcagcgcgaa	ggcctacagc	180
accgaagtg		aggagctgtg	ggccacccac	gctctgctgc	ccaccgacc	caacccccag	240
gagatctgcg		tgggagaaac	gacgcagaa	ttcaacatgt	ggaagaaacaa	catgttgtag	300
catgatcgac		aggacatcat	cagcctctgc	ggacgacgtg	tggagacctc	ctgtgaagctg	360
accccaactgt		gcgtgacctc	gcactcgacc	aacctgaaga	acgcaccaca	accgaagacc	420
agcaatcgta		aggagatgat	ccgcgcgtgc	atcaagaact	cgagctctaa	gctgaccacc	480
agcatccgca		acaagatgta	gaaggagctc	gctcctgtct	acaagctggt	ctgtgtgtccc	540
atcgacaaac		acacacacga	tcaacaagtg	atacactgca	acacacgctg	gatcaaccga	600
gctctgccca		agtgtagctt	ccgcgccatc	cccatccact	actgcgcccc	ccgcggcttc	660
gcacatctgc		agtgcaacga	caaagaattc	acggctgac	ccgcctcgca	caactgtgac	720
accgtgcagt		gcacccacga	ctatccggcc	gtgtgtagca	ccagctgact	gctgaagccc	780
acgtggcgct		agaggggcgt	ggtgatccg	acgagaaact	taccgcagaa	cgccaagcac	840
atcatcgtgc		gcattgaaga	gagctgggag	atcaactaca	ccgcgcccaa	caacaacacc	900
cgcaagagca		tcacctcgtc	ccccgcggcg	gcctttacgc	ccacggcgca	catcatcgcc	960
gacatctcgca		aggccctcgt	caacatcagc	ggcgagaagt	ggaacaacac	ctctgaagcag	1020
atcgtcgcca		acgtgcagcg	ccagttcggc	acaagaagca	tcgtgttcaa	gcagagcagc	1080
ggcgcgacc		ccgagatcgt	gtatgcacga	ttcaactcgc	gcggcgagtt	cttctactgc	1140
aacagcacc		actggtgtcaa	cagcatccgt	aacaacacgc	tcggccccaa	caacaccaac	1200
ggcacattca		ccctcgtcgt	ccgcatacgt	aggcccccct	acgccccccc	ctctccgcgc	1260
catgatccgt		gcagcagcaa	ctatcccgcg	ctgctgctgc	cccgcagaca	ggcgaagagat	1320
atcgacaaac		cccacagagt	cttccgcgcc	ggcgccgcgc	acatcgcgca	caactggcgc	1380
acgcagctgt		acaagtacaa	ggtgtgtga	atcgagcccc	tgggcgctgc	ccctcagac	1440
cccaagcgct		gcgtggtgca	gcgcgaaag	cgcgccgtgc	ccctggggcg	cgattctctg	1500
ggctctctgc		ggcgcgcgcg	cagcacatgt	ggcgcccgca	gcctcgacct	gacgctgcag	1560
ccgcgcacg		tgctgacgct	ctctgtgcga	cagcagacga	acctctggct	cgccatcact	1620
cgccagcagc		acctctgcga	gctgacgctg	tggggctatc	acgacgtgca	ggccgcgctgt	1680
gtccgcgtgc		acgcgtacct	gaaggacacg	gcctctgctg	gcattctgggg	ctctgcgcgc	1740
aagctcattc		gcacacccgc	cgtgcccctg	aaegccagct	ggagacaaca	gagctcagct	1800
catgatctga		acaacatgac	cttgatggag	tgggacgcg	agatgcacaa	taccacaac	1860
ctgatctaca		ccctgatcga	ggagagccag	aaccagcagc	agaagaacga	gcaggagctg	1920
ctggagctgt		acctatgggc	cagcctctgc	aactcgttgt	acatcagcaa	gtggctgtgtg	1980
tacatcaaga		ttcttatcat	gatctggggc	ggcctgtgtg	gctctgcagc	ctgtctcaac	2040
gtgtgagaca		tcgtgaacgc	cgtgcgcagc	ggctacaagc	ccctgagctt	ccagagaccgc	2100
ttcccgcgct		ccgcgcgcgc	gcagccgccc	aggggctatc	agggaagacc	ccggcagcgc	2160
gcagcgcagc		gcactgaccc	ctctgtgac	ggcctctgct	ccctgatctg	ggaagacctc	2220
ggcgacgtct		gctgtgtcac	ctaccaaccg	ctgcgcgacg	tgatctctg	ccgcgcgcgc	2280
atcgtggagc		tgctgtggcg	ccgcgctgtg	ggcgctgagc	agtactactg	caacctctct	2340
cagtaaatgc		tccagagctg	gaagaaacag	gcgcctgacc	tttgtagacg	catcgctcat	2400
gcgtgtggcg		agggcaccca	ccgatcatc	gaagtgcccc	acgcctactc	ccgcgctctc	2460
ctcacatctc		cccgcgccat	cccacgaagc	ttcaacaccc	ctctctctca	actcgac	2520

<210> 18  
<211> 2322  
<212> DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Leul22-Ser199;  
Arg426-Gly431

&lt;400&gt; 18

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagttcttcg ttctgcccag ccgcgtggag aagctgtggg tgaccgtgta ctacggcggtg 120
cccggtgtgga aggaggccac caccacctgt ttctgcgcga gcgcagccaa ggctctacgac 180
accggtgtgc accaactgtg ggccaccacc gctctgcgtg ccacgcagcc caacccccag 240
gagatcgtgc tggagaactg gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggcacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
ggcaacacag tgatcaccca ggccctgccc aaggtgagct tcagagccat ccccatccac 420
tactgcgccc ccgcgggctt cgccatcctg aagtgcacac acaagaagt caacggcagc 480
ggccctgtca ccaactgtag caccgtgtag tgcacccacc gcatccgccc cgtgtgtagc 540
accacgtgct tgcgtgaacg cagcctggcc gaggaggggc tgggtatccg cagcgaagac 600
ttcacgcaga accgcaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660
accgcgcccc acaacaacac ccgcaagagc atcaccatcg gccccggccc cgcttctcac 720
gccaccggcg acatcatcgg cgacatccgc caggcccact gcaacatcag cggcgagaag 780
tggaacaaca cctcgaagca gatcgtgacc aagctgcagg cccagttcgg caacaagacc 840
atcgtgttca agcagagcag cggcgccgac ccgagatcg tgatgcacag cttcaactgc 900
ggcgcgagat tcttctactg caacagcacc cagctgttca acagcaccgt gaacaacacc 960
atcgggcccc acaacaacca cggcaccatc accctgccct gcgcgaccaa gcatcatc 1020
aaccggcgcc cggcggaagg catgtacccc ccccccattc gcgcgcagat ccgctgcagc 1080
agcaaatcat ccggcctgct gctgaccgac gacggcgcca aggagatcag caacaccacc 1140
gagatcttcc gcccccggcg cggcgacatg cgcgacaact ggcgacgaga gctgtacaag 1200
tacaaggtgg tgaagatcga gcccttgggc gtggccccca ccaaggccaa gcgcgcggtg 1260
gtcgagcgcg agaagcgcg cgtgaccctg ggccgcatgt tcttgggctt ccttggcgcc 1320
gtcgagcaga ccatggcgcc ccgcagcctg accctgaccc tgacggcccc cagctgtgct 1380
agcggcatcg tgcagcaga gaacaacctg ctgcgcgcca tcagggccca gcagcactg 1440
ctgcagctga ccgtgtgggg catcaagcag ctgcagggccc cgctgctggc cgtggagcgc 1500
taccctgaag accagcagct gctgggcatc tggggctgca gcgcgaagct gatctgcacc 1560
accgcctgac cctggaacgc cagctggagc aacaagagcc tggaccagat ctggaacaac 1620
atgacctgga tggagtggga gcgcgagatc gacaaactca ccaactgat ctacaccctg 1680
atcgaggaga gccagaacca gcaggagaag aacgagcagg agctgctgga gctggacaag 1740
tggggcagcc tgtggaactg gttcgacatc agcaagtggc tgtgtgtacat caagatcttc 1800
atcatgatcg tggcgggcct ggtgggcctg cgcactgtgt tcaacctgct tcacactgct 1860
aaccgcgtgc gccagggtca cagccccctg agcttccaga cccgcttccc gcgccccccc 1920
ggccccgacc gccccgaggg catcgaggag gggggcgggc agcgcgaccg cgaccgcagc 1980
agccccctgg tgcacggcct gctggccctg atctgggagc agctgcgacg cctgtgctgt 2040
ttcagctacc accgctcgtg cgacctgac ctgatcgccc cccgcatcgt ggaagctgct 2100
gtcgcccgcg gctgggaagg cctgaagtag tggggcaacc tgcgtcagta ctggatccag 2160
gagctgaaga acagcgccgt gagcctgttc gacgcacat ccacgcgctg ggcgaggggc 2220
accgaccgca tcatcgaggt ggccccagcg atccggcgcc ccttctctga catccccccc 2280
cgatccgccc agggctctga ggcgcgctg ctgtaacctg ag 2322

```

&lt;210&gt; 19

&lt;211&gt; 2322

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Leul22-Ser199;  
Arg426-Lys432

&lt;400&gt; 19

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagttcttcg ttctgcccag ccgcgtggag aagctgtggg tgaccgtgta ctacggcggtg 120

```

```

cccggtgtgga aggaggccac caccaccctg ttctgcgcca gcgacgcca ggctctacgac 180
accgaggtgc acaacgtgtg ggcacccac gctcgctgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaa ttcaacatgt ggaagaacaa catgtgtggag 300
cagatgcacg aggcacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
ggcaacacg cgtatcacca ggccctgccc aaggtgagct tcgagcccat ccccatccac 420
tactggcgccc ccgcccgttt cgcacatcct aagtgcaacg acaagaagt caacggcagc 480
ggccctgtga ccaacgtgag caccgtgcag tgcaccacg gcatccgccc cgtgtgagc 540
accagcgtgc tgcgtgaacg cagcctggcc gaggagggtg tgggtatccg cagcgagaa 600
ttcacgcaca agcccaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660
accgcgcccc acaacaacac ccgcaagagc atcaccatcg gccccggccg gcctctctac 720
gccaccggcg acatcatcgg cgacatccgc caggcccact gcaacatcag cggcgagaag 780
tgggaacaaca cctgaagca gatcgtgacc aagctgcagg cccagttcgg caacaagacc 840
atcgtgttca agcagagcag cggcgccgac ccgagatcg tgatgcacag cttcaactgc 900
ggcgcgagat tcttctact caacagcacc cagctgttca acagcacctg gaacaacacc 960
atcgccccc acaacaacaa cggcaccatc accctgcctt gccgatcaa gcagatcctg 1020
aaccggcgcc gcaacaaggc catgtaccgc ccccccattc cggccagatc ccctgcgacg 1080
agcaacatca ccggcctgct gctgaccgc gacggcgcca aggagatcag caacaccacc 1140
gagatcttcc gcccccggcg cggcgacatg cgcgacaact ggcgacgaga gctgtacaa 1200
tacaagggtg tgaagatcga gccctgggc gtggcccaca ccaaggccaa ggcgcgctg 1260
gtgcagcgcg agaagcgcc cgtgaccctg ggcccatgt tcttgggctt cctggggcgc 1320
cgccggcaga ccatggggcg ccgcagcctg accctgaccg tgcaggcccg cagctctctg 1380
agcgccgatcg tgcagcaga gaacaacctg ctgcccacca tcgaggccca gcagcactg 1440
ctcgagctga ccgtgtgggg catcaagcag ctgcaggccc cgtgtgctgc cgtggagcgc 1500
ttcctgaagg accagacgt gctgggcctc tggggctgca cgggcaagct gatctgcaac 1560
accgcctgct cctggaacgc cagctgggag aacaagacc tggacagat ctggaacaac 1620
atgacatcga ttgagtgga gcgcgagatc gacacactga ccaactgat ctacactctc 1680
atcgaggaga gccgaacaa gcaggagaag aacgagcagg agctgctgga gctgagaac 1740
tggggcagcc tgtgtgaact gttcgacatc agcaagtgcc tgtgtgacat cagatatttc 1800
atcatgatcg tggcgccctt ggtggccctg cgcacatcgt taccgtgct gacgatcgtg 1860
aaccgcgtgc gccagggcta cagcccccct cgcctccaga cccgttccc cgcceccgcg 1920
ggcccccgcc gccccagggg catcgaggag gaggcgccg agcgcgacc cgacgcgacg 1980
agcccccctg gtcacggcct gctggccctg atctgggacg acctgcgac cctgtgcctg 2040
ttcagctacc accgcctcgt cgacctgatc ctgatcgccc cccgcacgt ggagctgctg 2100
ggccgcgcgc gctgggaggg cctgaagtac tggggcaacc tgcgcagta ctggatcgtg 2160
gagctgaaga acagcgcctg gacgctgttc gacgcccac ccatcgccgt ggccgagggc 2220
accgacgcga tcatcgaggt ggcccagcgc atcgccgcgc ccttctcgca catccccgcg 2280
cgcacccgcc agggcttcga gcgcgcctcg ctgtaactcg ag 2322

```

&lt;210&gt; 20

&lt;211&gt; 2322

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Leu122-Ser199;  
Trp427-Gly431

&lt;400&gt; 20

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgtgctg gctgtgtgga 60
cgagctcttc ttccccgcgc cgcgctggag aagctgtggg tgaccgtgta ctacggcgtg 120
ccggtgtgga aggaaggccac caccaccctg ttctgcgcca gcgacgcca ggctctacgac 180
accgaggtgc acaacgtgtg ggcacccac gctcgctgc ccaccgaccc caacccccag 240
gagatcgtgc tggagaacgt gaccgagaa ttcaacatgt ggaagaacaa catgtgtggag 300
cagatgcacg aggcacatcat cagcctgtgg gaccagagcc tgaagccctg cgtgaagctg 360
ggcaacacg cgtatcacca ggccctgccc aaggtgagct tcgagcccat ccccatccac 420
tactggcgccc ccgcccgttt cgcacatcct aagtgcaacg acaagaagt caacggcagc 480
ggccctgtga ccaacgtgag caccgtgcag tgcaccacg gcatccgccc cgtgtgagc 540
accagcgtgc tgcgtgaacg cagcctggcc gaggagggtg tgggtatccg cagcgagaa 600
ttcacgcaca agcccaagac catcatcgtg cagctgaagg agagcgtgga gatcaactgc 660

```

```

accgcgccca acaacaacac ccgcaagagc atcaccatcg gcccccgccg cgcctcttae 720
gccaccgcgg acatcatcgg cgacatccgc caggcccact gcaacatcag cggcgagaag 780
tggaacaaca cccctgaagca gatcgtgacc aagctgcagg cccagttcgg caacaagacc 840
atcgtgttca agcagagcag cggcgggcag cccagatcgt tgatgcacag cttcaactcg 900
ggcgggcgagt ttcttctact caacagcacc cagctgttca acagcacctg gaacaacacc 960
atcggcccca acaacaccaa cggcacccat accctgccct gccgcatcaa gcagatcatc 1020
aaccgctggg gcggcaaggc catgtacggc ccccccatcc gcggccagat cgtcgcagc 1080
agcaacatcaa cggcgctgct gctgaccggc gacggcgcca aggagatcag caacaccacc 1140
gagatcttcc gccccggcgg cggcgacatg cgcgacaact gcgcagcgga cgtgtacaag 1200
tacaagctgg tgaagatcga gcccttgggg gtggccccc ccaagggcaa gcgcgcgctg 1260
gtgcagcgcg agaagcgccg cgtgacctg ggcccatgt tcttgggctt cctggggcgc 1320
gccgagcaga ccatggggcg ccgcagcctg acctgaccg tgcaggcccg ccagctgctg 1380
agcggcatcg tgcagcagca gaacaacctg ctgcgcgcca tgcaggccca gcagcacctg 1440
ctgcagctga cctgtgtggg catcaagcag ctgcaggccc cgtgtgctgg cgtggagcgg 1500
tacctgaagg accagcagct cctgggcctc tggggctgca gcgcgaact gatctgcacc 1560
accgcgctgc cctggaacgc cagctggagg aacaagagcc tggaccagat ctggacaacc 1620
atgacctgga tggagtggga gcgcgagatc gacaactaca ccaacctgat ctacacctg 1680
atcgaggaga gccagaacca gcaggagaag aacagcagcg agctgctgga cgtggacaag 1740
tggggcagcc tgtggaactg gctcgacatc agcaagtggc tgtgtacat caagatcttc 1800
atcatgtatc tggcgggcct ggtgggcctg cgcactgtgt tcacogtgcg gacgatcgtg 1860
aaccgcgtgc gccagggcta cagccccctg agcttccaga ccgcgttccc gcgcccccgc 1920
ggccccgacc gccccgaggg catcgaagg agggcgcgcg agcgcgacc gcagccgagc 1980
agccccctgg tgcacggcct gctggccctg atctgggagc acctgcgacg cctgtgctgt 2040
ttcagctaac accgcctgcg cgacctgac ctgatcgcc ccgcgcatcg ggaactgctg 2100
ggccgcgcgg gctgggaggg cctgaagatc tggggcaacc tgcctgagta ctggatccag 2160
gagctgaaga acagcgccgt gagcctgttc gacgccaact ccactgcgct ggcgagggcg 2220
accgacgcca tcatcgaggt ggccccagcg atcgggcgcg ccttctctga catccccccc 2280
cgcatecgcc agggcttcca gcgcgcctg ctgtaacctg ag 2322

```

&lt;210&gt; 21

&lt;211&gt; 2310

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Lys121-Val200;

Asn425-Lys432

&lt;400&gt; 21

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
cgagttcttg ttctgcccag gcgcgtggag aagctgtggg tgacctgtta ctacggcgtg 120
ccgctgtgga aggaggccac caccaccctg ttctgcgcca gcgacgccaa ggccctacag 180
accgaggtgc acaacgtgtg ggccaccacc gcctgcgtgc ccaccgaccc caaccccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa ctggtggag 300
cagatgcacg aggaatcat catcctgtgg gaccagagcc tgaagccctg cgtgaaggcc 360
ccgctgatca cccagccctg ccccaaggtg agcttcgagc ccatcccatc ccactactgc 420
gcccccgcgg gcttcgcca cctgaagtgc aacgacaaga agttcaacgg cagcggcccc 480
tgaccacaag tgagcacctg gcagtgacc caccgcatcc gccccgtggt gagcacccga 540
ctgctgctga acggcagcct gcccgaggag ggctggtgta tccgcagcga gaacttcaac 600
gacaacgcga agaccatcat cgtgcagctg aaggagagcg tggagatcaa ctgcaccggc 660
cccaacaaca acaccggcaa gagcatcacc atcgcccccg gccgcgctt ctcgcgcaac 720
ggcgacatca tcggcgacat ccgccaaggc cactgcaaca tcagcgcgga gaagtgaac 780
aacaacctga agcagatcgt gaccaagctg caggcccagt tcggcaacaa gaccatcgtg 840
ttcaagcaga cgcgcggcgg cgaccccag atcgtgatgc acagcttcaa ctgcggcgcg 900
gagttctctt actgcaacag caccacagct ttcaacagca cctggaacaa caccatcgcg 960
cccaacaaca ccaacggcac catcacctg cctgcgccga tcaagcagat catcaacggc 1020
cccaaggcca tgtacgcccc ccccatcccg ggccagatcc gctgcagcag caacatcccc 1080
ggctcgtgct tgaccgcgca cggcggcgaag gagatcagca acaccacga gatcttcgg 1140
cccgcgcggg gcgacatgcg cgacaactgg cgcagcgagc tgtacaagta caaggtgggt 1200

```

```

aagatcgagc cctggggcgt gggccccacc aaggccaagc gccgcgtggt gcagcgcgag 1260
aagcgcgcgc tgaccctggg cgccatgttc ctgggcttcc tggggcgccgc cgccagcaacc 1320
atggggcgccc gcagcctgac cctgacacgtg caggccccgc agctgctgag cgcgcatcgtg 1380
cagcagcaga acaacctctgt gcgcgccatc gaggccccag agcactctgt cgactgagcc 1440
gtgtgggggca tcaagcagct gcaggccccg ctgctggcgt tggagcgcta cctgaaggac 1500
cagcagctgc tgggcatctg gggctgcagc ggcaagctga tctgaccacc cgcgctgccc 1560
tggaaacgcca ctgggagcaa caagagcctg gaccagatct gaaacaacat gacctggatg 1620
gagtgaggagc gggagatcga caactacacc aacctgatct acacctgat cgaggagagc 1680
cagaaaccagc aggaagaagaa cgagcaggag ctgctggagc tggacaagtg gggcagctgt 1740
tggaaactggt tgcacatcag caagtggctg tggatcacat agatcttcac catgatcgtg 1800
ggcgccctgg tgggcctgcg catcgtgttc acctgctgca gcatcgtgaa ccctgtgcgc 1860
cagggtctaca gcccctcgtg ctccagacc cgcttccccg ccccccgcg ccccgaccgc 1920
ccgaggggca tcgaggagga gggcgccgag cgcgaccgcg accgagcag cccctggtg 1980
cagcgcctgc tggccctgat ctgggacgac ctgcgcagcc tgtctcgtt cagctaccac 2040
cgctgcgcgc acctgatcct gatcgccgcc cgcctcgtg agctcgtggg ccgcccgcgc 2100
tgggagggccc tgaagtactg gggcaacctg ctgcagtact gcatccagga gctgaagaac 2160
agcgccgttg cctgttctga gcgcatcgcc atcgccgtgg ccgagggcac cgaccgcatc 2220
atcgaggtgg ccagcgcgat cggccgcgcc ttccctgaca tcccccgccg catccgcgag 2280
ggcttcgagc ggcctctgct gtaactcgag 2310

```

&lt;210&gt; 22

&lt;211&gt; 2298

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Val120-Ile201,  
Ile424-Ala433

&lt;400&gt; 22

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagttcttc tttegccag cgccgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccggttgga aggagggcac caccaccctg ttctgcgcca gcgacgcaa ggccatcagc 180
accgaggttg acaacgtgtg ggccaccac cctctgcgtg ccacgcagcc caaccgccag 240
gagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catggtggag 300
cagatgcacg aggaacatcat cagcctgtgg gaccagagcc tgaagccctg cgtggcgccg 360
atcacccagg cctgccccaa ggtgagcttc gagcccatcc ccatccacta ctgcgcccc 420
ggcggtcttc ccatctttaa gtgcaacgac aagaagtcca acggcagcgg cccctgcacc 480
aacgtgagca ccgtgcagtg caccacggcg atccgccccg tggtagacac ccagctgctg 540
ctgaacggca cgctggccga ggagggcggt gtgatccgca gcgagaactt caccgcacaa 600
gcgaagacca tcatcgtgca gctgaaggag agcgtggaga tcaactgcac ccgcccacac 660
aacaacaccc gcaagagcat caccatcgcg ccggcgccgc cttctacgc gaccggcgag 720
atcatcgccg acatccgcca ggcccactgc aacatcagcg gcgagaagtg gaacaacacc 780
ctgaagcaga tcgtgaccaa gctgcaggcc cagtttcgga acaagacatc cgtgttcaag 840
cagagcagcg cggcggaacc cgagatcgtg atgcacagct tcaactcgcc gcgcgagttc 900
ttctactgca acagcaccca gctgttcaac agcacctgga acaacacatc cgcccacac 960
aacaccaacg gcaccatcac cctgccctgc cgcatacagc agatcatcgg cggcgccatg 1020
tacgcccccc ccatccgcgg ccagatccgc tcgacgagca acatcacggc cctgctgctg 1080
accgcgagc ggcgcaagga gatcagcaac accacggaga tcttccgccc ccggggcgcc 1140
gacatcgccg acaactggcg cagcgaagctg tacaagtaca agtggtgtaa gatcgagccc 1200
ctggggctgg ccccaccaaa ggccaagcgc cgctgggtgc agccgagaa agccgcgctg 1260
acctcgggcg ccatgttctc gggcttccgt ggccgcgcgc gcagcacatc gggcgccgcg 1320
agctcgaccc tgacctgca ggcctgcgag ctgctgagcg gcatcgtgca gcagcagaa 1380
aacctgctgc ggcacatcga ggcccagcag cactctgctg agctgacgct gtgggggcatc 1440
aagcagctgc agggccgcgt gctggccgtg gagcgctacc tgaaggacca gcagctgctg 1500
ggcatctggg gctgcagcgg caagctgac tgcaccacgc ccgtgccctg gaagccagc 1560
tggagcaaca agagcctgga ccagatctgg aacaacatga cctgagtgga gtgggagcgc 1620
gagatcgaca actacaccaa cctgatctac acctgatcg aggagagcca gaccagcagc 1680
gagaagaacg agcaggagct ctgggagctg gacaagtggg ccagcctgtg gaactggttc 1740

```

gacatcgcca	agtggctgtg	gtacatcaag	atcttcac	tgatcgtggg	cggtcgtgtg	1800
ggcctgcgca	tcgtgttcac	cgtgctgagc	atcgtgaacc	cggtgcgcga	gggtcacagc	1860
ccctcgagct	tcacagaccg	cttcccgcgc	ccccgcggcc	ccgaccgccc	cgaggcgatc	1920
gaggaggagg	gcggcgcg	cgaccgcgac	cgcagcagcc	ccctggtgca	cggtcgtgtg	1980
gcctcgatct	gggacgacct	gcgcagcctg	tgccgtgtta	gctaccacgc	ctcgccgcag	2040
ctgatccctga	tgcgcgcgcg	catcgtggag	ctgctggggc	gcgcgcgctg	ggaggccctg	2100
aagtlactggg	gcaacctgct	gcagtlactgg	atccaggagc	tgaagaacag	cgctgtgagc	2160
ctgttcgcag	ccatcgccat	cgcctgtggc	gagggcacgc	acgcgatcat	cgaggtggcc	2220
cagcgcatcg	cgcgcgcctt	cctgcacatc	ccccgcgcga	tcgcgcaggg	cttcgagcgc	2280
gcctcgtctg	aactcgag					2298

&lt;210&gt; 23

&lt;211&gt; 2298

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence:

Val120-Ile201B; Ile424-Ala433

&lt;400&gt; 23

gaattcgcca	ccatggatgc	aatgaagaga	gggtctgtct	gtgtgtgtct	gctgtgtgga	60
gcagctcttc	tttcgccac	cgccgtggag	aagctgtggg	tgaccgtgta	ctacggcgctg	120
ccgtgtggga	aggaggccac	caccacccctg	tctgtgcgca	gcgacgcgca	ggctcacagc	180
accgagctgc	acaacgtgtg	ggccaccacc	gctcgtgtgc	ccaccgaccc	caaccgccag	240
gagatcgtgc	tgaggaaagt	gaccgagaaac	ttcaacatgt	ggaagaacaa	catggtgtgag	300
cagatgcgac	aggacatcat	cagcctgtgg	gaccagagcc	tgaagccctg	cgctgcgcgcg	360
atcacccagg	cgctccccaa	ggtgagcttc	gagcccatcc	ctatccacta	ctgcgcgcgcg	420
gccgctcttc	ccatctctgaa	gtgcaacgac	aagaagtcca	acggcagcgg	ccctgcaccc	480
aacgtgagca	ccgtgcagtg	caccacggcg	atccgcccgc	tggtgagcac	ccagctgtctg	540
ctgaacggca	gcctggccga	ggaggcgctg	gtgatccgca	gcgagaactc	caccgacaac	600
gccaaagacca	tcactcgtgca	gctgaaggag	agcgtggaga	tcaactgcac	ccgcccccaac	660
aacaacaccc	gcaagagcat	caccatcgcc	cccgcccgcg	ctttctacgc	caccgcgcag	720
atcatcgccg	acatccgcca	ggcccactgc	aacatcagcg	gcgagaagtg	gaacaacacc	780
ctgaagcaga	tgtgaccaa	gctgcaggcc	cagttccgca	acaagaccat	cgtgttcaag	840
cagagcgagc	gcggcgacc	cgagatcgtg	atgcacagct	tcaactcgcc	cggcgagttc	900
ttctactgca	acagcaccca	gctgttcaac	agcactcgga	acaacaccat	cggcccccaac	960
aacaccaaac	gcacatcac	cctgcctcgc	cgcatacaac	agatcatcgg	cggcgccatg	1020
tacgcccccc	ccatccggcg	ccagatccgc	tgacgacgca	acatcacccg	ctcgtctgtg	1080
accgcgcagc	gcggcaagga	gatcagcaac	accaccgaga	tcttccgcgc	cggcgccggc	1140
gacatgcgcg	acaactggcg	cagcgagctg	tacaagtaca	aggtggtgac	gactcagagc	1200
ctggcgctgg	ccccaccaca	ggccaagcgc	cgcgtggtgc	agcgcgagaa	cgcgccgctg	1260
acctctggcg	ccatgttctc	gggtctctct	ggcgccgcgc	cgacgacatc	ggcgccgcgc	1320
agcctgaccc	tgacgtgca	ggcccgccag	ctgctgagcg	gcactgtgca	cgacgagaa	1380
aacctgtctgc	gcgcacatga	ggcccgagcg	cacctgtctg	agctgacgtg	ctggggccatc	1440
aagcagctgc	aggcccgctg	gctggccgctg	gagcgctacc	tgaaggacca	gcagctgtctg	1500
ggcatctggg	gctgcagcgg	caagctgac	tgacacacg	ccgtgccccg	gaacgcagc	1560
tgatgcacaa	agagcctgga	ccagatctgg	aacaacatga	ccgtgagcga	gtgggagcgc	1620
ggatgcagca	actacaccaa	cctgatctac	acccctgatc	aggagagcca	gcacgcagag	1680
gagaagaacg	agcaggagct	gctggagctg	gacaagtggg	ccagcctgtg	gaactgtgtc	1740
gctgcagca	agtgctgtg	gtacatcaag	atcttcac	tgatcgtggg	cgccctggtg	1800
ggcctgcgca	tcgtgttcac	cgtgctgagc	atcgtgaacc	cggtgcgcga	gggtcacagc	1860
ccctcgagct	tcacagaccg	cttcccgcgc	ccccgcggcc	ccgaccgccc	caggggcatc	1920
gaggaggagg	gcggcgcgag	cgaccgcgac	cgcagcagcc	ccctggtgca	cggctcgtctg	1980
gcctcgatct	gggacgacct	gcgcagcctg	tgccgtgtta	gctaccacgc	ctcgccgcag	2040
ctgatccctga	tgcgcgcgcg	catcgtggag	ctgtgggccg	gcgcgcgctg	ggaggccctg	2100
aagtlactggg	gcaacctgct	gcagtlactgg	atccaggagc	tgaagaacag	cgctgtgagc	2160
ctgttcgcag	ccatcgccat	cgcctgtggc	gagggcacgc	acgcgatcat	cgaggtggcc	2220
cagcgcatcg	gcgcgcctt	cctgcacatc	ccccgcgcga	tcgcgcaggg	cttcgagcgc	2280



gccccgtgctg aactcgag

2298

&lt;210&gt; 24

&lt;211&gt; 2298

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Val120-Thr202;  
11e424-Ala433

&lt;400&gt; 24

```

gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagtcttcc ttctgcccag cggcgtggag aagctgtggg tgaccgtgta ctacggcgtg 120
cccgctgtgga aggagggcac caccaccctg ttctgcccga gcgacgcaa ggccctacgac 180
accgaggtgac acaacgtgtg gggcaccacc gctcgctgtg ccaccgaccc caaccgccac 240
cagatcgtgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catgggtggag 300
cagatgcacg aggcacatcat cagcctgtgtg gaccagagcc tgaagccctg cgtgggctgg 360
ggcaaccagg cctgcccmaa ggtgagcttc gagccatcc ccattcacta ctgcccctcc 420
gcccgtcttc ccattcctgaa gtgcaacgac aagaagtcca acggcgcggg cccctgcacc 480
aacgtgagca cegtgcagtg caccacgggc atccgcccgc tggtagacac caagctgctg 540
ctgaaccgca gcttgccgga ggagggcgtg gtgatccgca gcgagaactt caccgacaa 600
ggcaagaccca tcactgtgca gctgaaggag agcgtggaga tcaactgcac ccgcccacac 660
aacaacacccc gcaagagcat caccatcggc cccggccgcg ccttctacgc caccggcgac 720
atcatcggcg acatccgcca gggccactgc aacatcagcg gcgagaaagt gaacaacacc 780
ctgaagcaga tegtaccmaa gctgcaggcc cagttcggca acaagaccat cgtgttcaag 840
cagagcagcg gcggcgaccc cgagatcgtg atgcacagct tcaactcgcg cggcgagttc 900
ttctactgca acagcaccga cctgttcaac agcactgga acaacaccat cggcccacac 960
taacaccaac gcaccatcac cctgcccctg cgcacatcag agatcactcg cggcgccact 1020
tgcgcccacc ccattccggg ccagatccgc tgcagcagca acatcaccgg cctgctgctg 1080
accgcgacg gcggcaagga gatcagcaac accaccgaga tcttcgcccc cggcgccggc 1140
gacatgcgcg acaactggcg cagcagctg tacaagtaca agtggtgaa gatcgagccc 1200
ctggcgctgg ccccccacaa ggccaaagcg cgcgtggtgc agcgcgagaa gcgcccgtg 1260
accctggggc ccatgttctt gggcttctgt ggcgcgcgcg gcagcaccaat ggcgcgccgc 1320
agcctgaccc tgacgtgca ggcgcgccag ctgctgagcg gcatcgtgca gcagcagaac 1380
aaactcgtgc gcgccatcga ggcgcagcag cactgctgc agctgacgtg gtggggcattc 1440
aagcagctgc aggcgcgctg gctggcgtg gagcgtacc tgaaggacca gcagctgctg 1500
ggcatctggg gctgcagcgg caagctgata tgcaccaacc cctgcccctg gaacgccagc 1560
tggagcaaca agnagcctgga ccagatctgg aacaacatga cctggatgga gtgggagcgc 1620
gagatcgaca actacaccaa cctgatctac accctgatcg agggagacca gaaccagcag 1680
gagaagaacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactgtgtc 1740
gacatcagca agtggtctgtg gtacatcaag atcttcatca tgatcgtggg cggcctgtgtg 1800
ggcctgcgca tctgtttcac cgtgctgagc atcgtgaacc gcgtgcgcca gggctacagc 1860
cccctgagct tccagaccgg cttcccgcgc ccccgcgccc ccagcgccc ccagggcattc 1920
ggagagaggg gcggcgagcg cgacgcgac cctgggtgca cctgctgctg gtaggctgctg 1980
ggcctgatct gggacgacct gcgcagcctg tgcctgttca gctaccacgc cctgcgcgac 2040
ctgatcctga tgcgcgcccg catcgtggag ctgctggggc gcgcggctg gtaggctcctg 2100
aagtactggg gcaacctgct gcagtactgg atccaggagc tgaagaacag ccgcgtgagc 2160
ctgttcagcg ccatcgccat cggcgtggcc gagggcaccc accgcatcat cgaggtggcc 2220
cagcgcacgt gcgcgcgctt cctgcacatc cccgcgcgca tccgccaggg ctttcgagcg 2280
gcctgctgtg aactcgag

```

2298

&lt;210&gt; 25

&lt;211&gt; 2358

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Val127-Asn195

```

<400> 25
gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagttcttc ttctgccacc gcgcgtggag aagctgtggg tgaccgtgta ctacggcggtg 120
cccggtgtgga aggaggccac caccaccttg ttctgcgcca gcgacgccaa ggccctacgac 180
accgaggtgtg acaacgtgtg ggccaccacac gcttgcgtgc ccaccgacct caaccctccag 240
gagatcgttgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catgggtggag 300
cagatgcacg aggcacatcat cagcctgttg gaccagagcc tgaagccctg cgtgaagctg 360
acccccctgt gctgtggggc agggaaactgc aacaccagcg tgatcaccca ggccctgccc 420
aaggtgagct tgcagcccc ccccatccac tactgcccc ccgcccgtct cgcacatcgt 480
aagtgcacag acaagaagtt caacggcgag ggcctctgca ccaacgtgag ccagctcgag 540
tgcaccacag cctccgcgcc cgtggtgagc accacgtgcg tgctgaacgg cagcctggcc 600
gaggaggggc tgggtgatcc cagcgagaac ttaccgcgca acgccaagac catcatcgtg 660
cagctgaagg agagcgtgga gatcaactgc acccgcccc acaacaacac ccgcaagagc 720
atcaccatcg gccccggccc cgccttctac gccaccggcg acatcatcgg cgacatccgc 780
caggcccact gcaacatcag cggcgagaag tggaaacaac cctgaagca gatcgtgacc 840
aagctgcagg cccagttccg caacaagacc atcgtgttca agcagagcag ccgcccggac 900
cccgagatcg tgatgcacag ctccaactgc ggcggcgagt tcttctactg caacagcacc 960
cagctgttca acagcacctg gaacaacacc atcgcccca acaacaacca cggcaccatc 1020
accctgccct gccgcataca gcagatcac aacgcgtggc aggaggtggg caaggccatg 1080
tacgcccccc ccatccgcgg ccagatccgc tgacgcagca acatcacggg cctgctgctg 1140
accgcgcagc ggcgaagga gatcagcaac accaccgaga tcttcgcccc cggcgccggc 1200
gacatcgccc acaactggcg cagcgagctg tacaaagtaca aggtggtgaa gctcgagccc 1260
ctgggctgtg cccccccaaa ggccaagcgc cgtgtgtgtg agcgcgagaa gcgcccgtg 1320
accctggggc ccatgttctc gggcttctcg ggcgcgcggc gcagcaccat gggcgcccgc 1380
agcctgaccc tgaccgtgca ggcgcgcgag ctgctgagcg gcatcgtgca gcagcagaac 1440
aacctgctgc gcgccatcga ggcgccagc cactgtgtgc agctgacgt gtgggggcatc 1500
aagcagctgc agccccgcgt gctggccgtg gagcgtacc tgaaggacca gcagctgctg 1560
ggcatctggg gctgcagcgg caagctgac tgacacacgc ccgtgccctg gaacgcgagc 1620
tggagcaaca agagccttga ccagatctgg aacaacatga cctggatgga gtgggagcgc 1680
gacatcagca actacaccaa cctgatctac accctgatgc aggagagcca ccaccagcag 1740
gagaaagacg agcaggagct gctggagctg gacaagtggg ccagcctgtg gaactggttc 1800
gacatcagca agtggtgtg gtacatcaag atcttcatca tgatcgtggg ccgctctggtg 1860
ggctgcgcca tctgttacc cgtgctgagc atcgtgaacc cgtgcgcca gggtctacagc 1920
ccctgagct tccagaccgc cttccccgcc ccccgcgccc ccgaccgccc caggggcatc 1980
gaggagaggg ggcgcgagcg gcagcgcgac cgcagcagcc cctgtgtgca cggctctgctg 2040
gacctgatct gggacgacct gcgcagcctg tgctgttcta gctaccaccg cctgcgcgac 2100
ctgatctgta tgcgcgcgcc catcgtggag ctgctggggc gcgcgcgctg ggaaggcctg 2160
aagttactgg gcaacctgct gcagttactg atccaggagc tgaagaacag ccgcgtgagc 2220
ctgttcgacg ccatcgccat gcctgtggcc gaggggcagc accgcacatc caggttggcc 2280
cagcgcacg gcgcgcctt cctgcacatc cccgcgcgca tccgcagagg cttcgcagcg 2340
gcctctgctg aactcag 2358

```

<210> 26  
 <211> 2352  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Vall127-Asn195;  
 Arg426-Gly431

```

<400> 26
gaattcgcca ccatggatgc aatgaagaga gggctctgct gtgtgctgct gctgtgtgga 60
gcagttcttc ttctgccacc gcgcgtggag aagctgtggg tgaccgtgta ctacggcggtg 120
cccggtgtgga aggaggccac caccaccttg ttctgcgcca gcgacgccaa ggccctacgac 180
accgaggtgtg acaacgtgtg ggccaccacac gcttgcgtgc ccaccgacct caaccctccag 240
gagatcgttgc tggagaacgt gaccgagaac ttcaacatgt ggaagaacaa catgggtggag 300
cagatgcacg aggcacatcat cagcctgttg gaccagagcc tgaagccctg cgtgaagctg 360
acccccctgt gctgtggggc agggaaactgc aacaccagcg tgatcaccca ggccctgccc 420

```

```

aaggtagctg tgcagcccat ccccatccac tactgcgccc ccgcccgtt cgcacatcct 480
aagtgcacac acaagaagtt caacgcgcgc ggcacctgca ccaactgag caccgtgcag 540
tgcacccacg gcatcccgccc cgtggtgagc acccagctgc tgcctgaacgg cagcctgggc 600
gaggaagggc tgggtatccg cagcgagaac ttaccgcaca acgccaagac catcatcgtg 660
cagctgaagg agagcgtgga gatcaactgc acccgcccca acaacaacac ccgcaagagc 720
atcaccatcg gccccggcg cgccttctac gccaccggcg acatcatcgg cgacatccgc 780
caggcccact gcaacatcag cggcgagaag tggacaaca cctgaagca gatcgtgacc 840
aagctgcagg cccagttcgg caacaagacc atcgtgttca agcagagcag cggcgcgacc 900
cccgagatcg tgatgcacag cttcaactgc ggcggcgagt tcttctactg caacagcacc 960
cagctgttca acagcacctg gaacaacacc atcgccccca acaacaccaa cggcaccatc 1020
acctgccccg gcgcgcatcaa gcagatcatc aacgcggcg gcggcaaggc catgtacgcc 1080
cccccatcc gcggccagat ccgctgcgc agcaacatca ccgcccctgt gctgacccgc 1140
gacggcgcca aggagatcag caacaccacc gagatcttcc gccccggggg cggcgacatg 1200
cgcgacaact ggcgacgcga gctgtacaag tacaagtggt gaagatcga gccctgggc 1260
gtggccccc acaaggccaa gcgcgcgctg gtgcagcgcg agaagcgcg cgtgacctg 1320
ggcgccatgt tectgggctt cctggcgccg gccggcagca ccatggggcg ccgcagcctg 1380
acctgacgcg tgcaggcccg ccagctgctg agcggcatcg tgcagcagca gaacaacctg 1440
ctgcgcgcga tgcaggccca gcagcacctg ctgcagctga ccgtgtgggg catcaagcag 1500
ctgcaggccc gcgtgctggc cgtggagcgc tacctgaagg accagcagct gctgggcacg 1560
tggggctgca gcggcaagct gatctgcacc accgcgctgc cctggaacgc cagctggagc 1620
aacaagagcc tggaccagat ctggaacaac atgacctgga tggagtggga gcgcgagatc 1680
gacaactaca ccaacctgat ctacacctg atcgaggaga gccagaacca gcaggagaag 1740
aacgagcagg agctgctgga gctggacaag tgggccagcc tgtggaactg gttcgacatc 1800
agcaagtggc tgtggtacat caagatcttc atcatgatcg tgggcggcct ggtgggcctg 1860
cgcatcgtgt tcacgtgctg gagcatcgtg aaccgcgtgc gccagggcta cagccccctg 1920
agcttcacga cccgcttccc gcccccccg gcccccgacc gccccgagg catcgaggag 1980
gagggcgcg agcgcgacc cgacgcgcgc agccccctg tgcacggcct gctggccctg 2040
atctgggacg acctgcgcag cctgtgcctg ttacgtacc accgcctgcg cgacctgac 2100
ctgatcgcg cccgcatcgt ggagctgctg ggccgcgcg cctgggaggc cctgaagtac 2160
tggggcaacc tgcctgagta ctggatccag gagctgaaga acagcgcggt gagcctgttc 2220
gacgccatcg ccatcgccgt ggccgagggc accgaacgca tcatcgaggt ggcacagcg 2280
atcggcgcg ccttctgca catccccgc cgcattccgc agggcttcga gcgcgcctg 2340
ctgtaactcg ag 2352

```